

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 19:05:32 ; Search time 77 Seconds
(without alignments)
8.246 Million cell updates/sec

Title: US-09-813-341-10
Perfect score: 20
Sequence: 1 GSGS 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03.*

1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	4	19	Mutant IJ loop of
2	20	100.0	4	20	HCV NS4A-NS3 Compl
3	20	100.0	4	22	Linker sequence #1
4	20	100.0	4	22	Linker peptide #2
5	20	100.0	4	22	Peptide spacer. U
6	20	100.0	4	22	Peptide associated
7	20	100.0	4	23	Tetrapeptide linker
8	20	100.0	4	23	HIV-1 Tat protein
9	20	100.0	4	23	Fd-flexible linker

10	20	100.0	4	24	ABG73964	Synthetic peptide
11	20	100.0	5	19	AAW56059	Mutant CD loop of
12	20	100.0	5	21	ABO7386	Human alpha d meta
13	20	100.0	5	23	ABJ0541	Target fusion pep
14	20	100.0	5	24	AAO26595	Fusion protein rel
15	20	100.0	6	19	AAW56062	Mutant FG loop of
16	20	100.0	6	20	AAV15209	Optional sequence
17	20	100.0	6	22	AAU04632	Linker sequence #2
18	20	100.0	6	22	AAO04504	Linker peptide #3
19	20	100.0	6	22	AAO75115	Linker peptide seq
20	20	100.0	6	23	ABP56401	Peptide linker #1
21	20	100.0	6	23	AAE16123	Phosphopeptide #1
22	20	100.0	7	19	AAW70511	Escherichia coli S
23	20	100.0	7	20	AAV40714	S3 derivative #11
24	20	100.0	7	20	AAV40715	S3 derivative #12
25	20	100.0	7	20	AAV40716	S3 derivative #13
26	20	100.0	7	21	AAE30053	Scaffold protein S
27	20	100.0	7	21	AAE30054	Scaffold protein S
28	20	100.0	7	21	AAE30055	Scaffold protein S
29	20	100.0	7	23	AAU97979	Peptide linker for
30	20	100.0	8	14	AAE38551	Peptide #1 for the
31	20	100.0	8	17	AAW05710	Minimal motif #6..
32	20	100.0	8	19	AAW79115	Gly-ala polymer of
33	20	100.0	8	21	AAE28998	Peptide associated
34	20	100.0	8	22	AAU04616	Linker sequence us
35	20	100.0	8	22	AAE04488	Linker peptide #1
36	20	100.0	8	24	AAE30841	Gly-Ser tag peptid
37	20	100.0	9	14	AAE42528	Leukocyte-binding
38	20	100.0	9	17	AAW49247	Human leucocyte an
39	20	100.0	9	21	AAE15365	Modified human LHR
40	20	100.0	9	21	AAE23369	Human trophinin ep
41	20	100.0	9	22	AAE09530	Human mucin-1 (MUC
42	20	100.0	9	22	AAE60933	Spacer peptide. S
43	20	100.0	9	22	AAE08093	Modified humanised
44	20	100.0	9	22	AAU03378	Trar/GS/BIAD fusio
45	20	100.0	9	22	AAU02650	CDR region of anti

ALIGNMENTS

RESULT 1

AAW56063
ID AAW56063 standard; peptide; 4 AA.

XX AC AAW56063;

XX DT 29-JUL-1998 (first entry)

XX DE Mutant IJ loop of the Ad5 fiber knob.

XX KW Chimeric; adenovirus; fiber protein; binding; targeting; coat protein;
KW constrained peptide motif; gene therapy; cancer; heart disease;
KW autoimmune disorder.

XX OS Synthetic.

XX OS Mastadenovirus.

XX PN WO9807865-A1.

XX PD 26-FEB-1998.

XX PF 21-AUG-1997; 97WO-US14719.

XX PR 21-AUG-1996; 96US-0701124.

XX PA (GENV-) GENVEC INC.

XX PI Kovesdi I, Roelvink PW, Wickham TJ;

XX DR WPI; 1998-169169/15.

XX PT Chimeric adenovirus fibre proteins - containing non-native amino

PT acid sequence to provide for binding and entry into cells,
 PT especially for gene therapy

XX Example 1; Page 46; 124pp; English.

XX The present sequence represents a mutant IJ loop of the Ad5 fiber
 CC knob, which is used in an example from the present invention. The
 CC present invention describes a chimeric adenovirus fibre protein (APP)
 CC containing a constrained non-native amino acid sequence. The non-native
 CC amino acid sequence allows the chimeric fibre (or a vector comprising
 CC the chimeric fibre) to more efficiently bind to and enter cells. The
 CC products can be used for gene therapy, for treating cancer, e.g.
 CC melanoma, glioma and lung cancers as well as genetic disorders, e.g.
 CC cystic fibrosis, haemophilia and muscular dystrophy as well as
 CC pathogenic infections, e.g. HIV, tuberculosis and hepatitis and also for
 CC heart disease, to e.g. prevent restenosis following angioplasty or to
 CC promote angiogenesis to reperfuse necrotic tissue, and in autoimmune
 CC disorders, e.g. Crohn's disease, colitis, rheumatoid arthritis, and
 CC Alzheimer's disease.

XX SQ Sequence 4 AA;

Query Match 100.0%; Score 20; DB 19; Length 4;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GSGS 4
 Db 1 GSGS 4

RESULT 2

AA117875
 ID AAY17875 standard; peptide; 4 AA.

AC AAY17875;

DT 07-SEP-1999 (first entry)

DE HCV NS4A-NS3 complex linker #1.

XX HCV; hepatitis C virus; single chain recombinant complex; linker;
 KW NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease;
 KW hydrophobic domain; covalent complex; detection; inhibitor.

OS Synthetic.

OS Hepatitis C virus.

XX WO9928482-A2.

XX 10-JUN-1999.

XX 24-NOV-1998; 98WO-US24528.

XX 28-JUL-1998; 98US-0094331.

XX 28-NOV-1997; 97US-0067315.

XX (SCHE) SCHERING CORP.

XX Malcolm BA, Taremi SS, Weber PC, Yao N;

XX WPI; 1999-385385/32.

XX New hepatitis C virus covalent complexes

XX Claim 5; Page 107; 211pp; English.

XX The present invention describes a covalent hepatitis C virus (HCV)
 CC NS4A-NS3 complex comprising a central hydrophobic domain of native HCV
 CC NS4A peptide, a linker, and an HCV NS3 serine protease domain, where the
 CC hydrophobic domain of native HCV NS4A peptide is tethered by the linker
 CC to the amino terminus of the HCV NS3 protease domain. The present
 CC sequence represents a specifically claimed linker for use in the above

CC complex. The covalent NS4A-NS3 complexes are useful for structural
 CC determination and determination of mode of binding of HCV inhibitors by
 CC NMR spectroscopy. They can also be used for detecting inhibitors of the
 CC protease activity, the helicase activity and the ATPase activity of NS3.
 CC The covalent NS4A-NS3 complexes are more soluble, stable and active than
 CC the non-covalent protease-peptide complexes previously available.

XX SQ Sequence 4 AA;

Query Match 100.0%; Score 20; DB 20; Length 4;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GSGS 4
 Db 1 GSGS 4

RESULT 3

AAU04631

ID AAU04631 standard; Protein; 4 AA.

AC AAU04631;

DT 23-OCT-2001 (first entry)

XX Linker sequence #1 used to make single chain gonadotropin analogues.

XX Human; glycoprotein hormone; infertility; in vivo fertilisation;

KW single chain gonadotropin.

OS Synthetic.

XX US6242580-B1.

XX 05-JUN-2001.

XX 31-MAR-1999; 99US-0282357.

XX 25-AUG-1997; 97US-0918288.

XX 18-FEB-1994; 94US-0193882.

XX 12-AUG-1994; 94US-0289396.

XX 23-SEP-1994; 94US-0310590.

XX 04-NOV-1994; 94US-0334628.

XX 07-DEC-1994; 94US-0351591.

XX 07-JUN-1995; 95US-0475049.

XX 09-MAY-1997; 97US-0853524.

XX (UNIW) UNIV WASHINGTON.

XX Boime I, Moyle WR;

XX WPI; 2001-424301/45.

XX New single chain forms of the glycoprotein hormone quartet useful for
 PT generating antibodies specifically immunoreactive with the new
 PT compounds, in treating infertility, or as aids for in vivo
 PT fertilization techniques

XX Example 19; Column 36; 86pp; English.

XX The sequence represents the amino acid sequence of the linker #1 used to
 CC make single chain gonadotropin analogues. The glycoprotein
 CC hormone analogues are useful for generating antibodies specifically
 CC immunoreactive with new compounds, as substitutes for the heterodimeric
 CC forms of the hormones, in the treatment of infertility, as aids for in
 CC vivo fertilization techniques, and other therapeutic methods associated
 CC with the native hormone. The single chain protein is also useful as a
 CC reagent in a manner similar to the heterodimer, as a diagnostic tool to
 CC detect the presence of antibodies with respect to the native proteins in
 CC the biological samples, as a control reagent in assay kits for assessing
 CC the levels of these hormones in various samples, and in detecting and
 CC purifying receptors to which the native hormones bind. The single chain

CC forms of the heterodimers or homodimers have the following advantages
 CC over their dimeric forms: they are more stable, problems of recombinant
 CC production are reduced since only a single gene is needed to transcribe,
 CC translate and process, provide an alternate form thus permitting fine
 CC tuning of activity levels and of in vivo half lives. Single chain forms
 CC are unique starting materials for identifying truncated forms with the
 CC activity of the dimer. The linkage between the subunits permits the
 CC protein to be engineered without disturbing the overall folding of the
 CC protein.
 XX SQ Sequence 4 AA;
 Query Match 100.0%; Score 20; DB 22; Length 4;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GSGS 4
 Db 1 GSGS 4
 RESULT 4
 ID AAE04503 standard; peptide; 4 AA.
 XX AAE04503;
 XX 04-SEP-2001 (first entry)
 DT Linker peptide #2.
 DE Human: single chain gonadotropin analog; anti-infertility; drug;
 KW peptide therapy; luteinising hormone; LH; follicle stimulating hormone;
 KW FSH; thyroid stimulating hormone; TSH; chorionic gonadotropin; CG;
 KW glycoprotein; infertility; fusion protein; linker.
 XX Synthetic.
 OS US6238890-B1.
 XX 29-MAY-2001.
 PD 25-AUG-1997; 97US-0918288.
 PF 18-FEB-1994; 94US-0199382.
 PR 12-AUG-1994; 94US-0289396.
 PR 22-SEP-1994; 94US-0310590.
 PR 04-NOV-1994; 94US-0334628.
 PR 07-DEC-1994; 94US-0351591.
 PR 07-JUN-1995; 95US-0475049.
 PR 09-MAY-1997; 97US-0853524.
 XX (UNIW) UNIV WASHINGTON.
 PA Boime I, Moyle WR;
 PI WPI; 2001-366474/38.
 XX New DNA or RNA encoding single chain protein useful in treating
 PT infertility, as aids in vitro fertilization techniques, or other
 PT therapeutic methods associated with the native hormones -
 PT Example 19; Column 35; 87pp; English.
 PS The invention relates to human single chain forms of the glycoprotein
 CC hormone quartet which is an agonist or antagonist of luteinising hormone
 CC (LH), follicle stimulating hormone (FSH), thyroid stimulating hormone
 CC (TSH) or chorionic gonadotropin (CG). All these hormones are heterodimers
 CC having identical alpha subunits and differing beta subunits. The agonist
 CC forms of single chain hormones are used in treating infertility, as aids
 CC in vitro fertilization techniques, and other therapeutic methods
 CC associated with the native hormones. The single chain hormones are useful
 CC as reagents in a manner similar to heterodimers, as diagnostic tools to

CC detect the presence of antibodies with respect to the native proteins in
 CC biological samples, as control reagents in assay kits for assessing the
 CC levels of these hormones in various samples, in detecting and purifying
 CC receptors to which the native hormones bind. The single chain hormones
 CC are also used in affinity chromatographic preparation of receptors or
 CC antihormone antibodies. They are used as purification tools for
 CC isolation of subsequent preparations of these materials and to monitor
 CC levels of single chain hormones administered as drugs. The single chain
 CC glycoproteins are used to generate antibodies specifically immunoreactive
 CC with these new compounds, as substitutes for the heterodimeric forms of
 CC hormones. The present sequence is linker peptide which is used for
 CC constructing single chain gonadotropin analogs related to the invention.
 CC Analog fusion proteins serves as useful starting compounds for template
 CC directed vaccine design and for the development of hormone-specific
 CC vaccines for use in humans.
 XX SQ Sequence 4 AA;
 Query Match 100.0%; Score 20; DB 22; Length 4;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GSGS 4
 Db 1 GSGS 4

RESULT 5
 AAB87966
 ID AAB87966 standard; peptide; 4 AA.
 XX AAB87966;
 AC AAB87966;
 XX 17-MAY-2001 (first entry)
 DT Peptide spacer.
 DE Tat; HIV-1; human immunodeficiency virus; AIDS; viremia.
 KW Unidentified.
 XX US6193981-B1.
 PN 27-FEB-2001.
 PD 10-JUL-1998; 98US-0113921.
 PF 11-JUL-1997; 97US-0893853.
 XX (THYM-) THYMON LLC.
 PA Goldstein G;
 PI WPI; 2001-243400/25.
 DR New compositions comprising at least two variants of a Tat protein
 PT coupled to a carrier protein, useful for impairing multiplication of
 PT human immunodeficiency virus 1 and in eliciting anti-Tat antibodies -
 PT Disclosure; Column 23; 63pp; English.
 PS The present invention relates to a composition with at least two
 CC variants of Tat protein coupled to a carrier protein. The invention is
 CC useful for inhibiting the multiplication of HIV-1 virus in infected
 CC patients with viremia, symptomatic or asymptomatic and for
 CC attenuating HIV-1 multiplication during primary infection in
 CC previously uninfected subjects.
 XX SQ Sequence 4 AA;
 Query Match 100.0%; Score 20; DB 22; Length 4;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

XX 04-DEC-2001.
PD
XX
XX 25-JUN-1999; 99US-0344456.
XX
XX 25-JUN-1999; 99US-0344456.
XX
XX (SCHE ) SCHERING CORP.
XX
XX Hong Z, Lai VCH, Lau JYN;
XX
XX WPI; 2002-121103/16.
XX
XX Nucleic acid construct encoding chimeric Hepatitis C Virus (HCV)
XX pestivirus genome where the Npro protease gene is replaced with NS3
XX protease gene, useful for in vivo screening of compounds which inhibit
XX HCV infection
XX
XX Example 1; Column 13; 20pp; English.
XX
XX The present invention relates to a nucleic acid construct encoding a
XX chimeric Hepatitis C virus (HCV)-pestivirus genome. The construct
XX comprises a pestivirus genome where a Npro pestivirus protease gene is
XX replaced with a gene encoding a functional HCV NS3 protease. Furthermore,
XX each junction site recognised by the Npro protease is replaced with a
XX junction site recognised by the HCV NS3 protease. The construct is useful
XX for screening compounds that inhibit HCV in vivo by inhibiting HCV
XX protease, where screening may be in cell culture or in an animal model.
XX The present peptide linker was used to illustrate the present invention.
XX
XX Sequence 4 AA;
XX
XX Query Match 100.0%; Score 20; DB 23; Length 4;
XX Best local similarity 100.0%; Pred. No. 9.3e+05;
XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX OY 1 GSGS 4
XX ||||
XX 1 GSGS 4
XX
XX Db
XX
XX RESULT 8
XX AAU11093
XX
XX IID AAU11093 standard; peptide; 4 AA.
XX
XX AAC
XX AAU11093;
XX
XX 14-FEB-2002 (first entry)
XX
XX HIV-1 Tat protein epitope C-terminal spacer peptide #2.
XX
XX HIV; human immunodeficiency virus; Tat; transactivating protein;
XX epitope 1; immunogen; vaccine; antiviral; antibody; spacer.
XX
XX Synthetic.
XX
XX WO200182944-A1.
XX
XX 08-NOV-2001.
XX
XX 20-APR-2001; 2001WO-US13031.
XX
XX 28-APR-2000; 2000US-0561366.
XX
XX (THYM-) THYMOM LLC.
XX
XX Goldstein G;
XX
XX WPI; 2002-055423/07.
XX
XX New composition, useful for impairing human immunodeficiency virus-1
XX multiplication, comprises at least 2 variants of (poly)peptide or
XX antibody immunospecific for (poly)peptide -

```

XX Disclosure; Page 35; 84pp; English.

XX The invention relates to a composition comprising at least 2

XX variants of a polypeptide, the polypeptide being based on an epitope

XX sequence from the HIV (human immunodeficiency virus) Tat

XX (transactivating) protein. Also included are antibodies raised

XX against the polypeptides, nucleic acids encoding the polypeptides,

XX methods of determining the antibody titre in a subject immunised with the

XX polypeptides and microorganisms trans formed with the polynucleotides

XX which express the polypeptides. The compositions, antibodies, synthetic

XX genes and molecules and microorganisms are useful for impairing the

XX multiplication of HIV-1. The composition is also useful for assessing the

XX immune status of vaccinated patients by determining the presence and/or

XX titre of antibodies induced by immunisation to a Tat immunogen.

XX An advantage of the composition is the small number of immunogens

XX required for inclusion into the composition to cross react with greater

XX than 95-99% of known Tat protein variants of HIV-1 of the common B

XX subtype. The prime recognized Epitope I immunogenic composition

XX containing the 2 required prime recognized Epitope I amino acid

XX sequences as well as the additional Epitope I sequences cross reacts with

XX 95% Tat proteins of HIV-1 of the common B subtype as well as with all 56

XX Tat protein sequences from less frequent non-B subtypes of HIV-1.

XX Therefore a single composition may be usefully employed in protecting

XX against or treating infection, caused by the vast majority of HIV-1

XX strains that can be encountered. New desirable immunogens from new

XX occurring HIV-1 strains may be easily identified using the new methods.

XX This enables the compositions to be useful prophylactically against any

XX new strains of HIV-1 identified in the future. The present sequence

XX is a spacer peptide used to link a Tat epitope based sequence of

XX the invention to a solid support when included in a kit for assaying the

XX titre of the antibodies.

XX Sequence 4 AA;

SQ Query Match 100.0%; Score 20; DB 23; Length 4;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSGS 4
|||

Db 1 GSGS 4

RESULT 9

AG78441

XX ID AAG78441 standard; protein; 4 AA.

XX AC AAG78441;

XX DT 12-APR-2002 (first entry)

XX DE Pd-flexible linker peptide.

XX Antibody; antigen; immunoglobulin; ADCC; CDC; anti-globin response;

XX antibody dependant cell mediated cytotoxicity;

XX complement dependant cytotoxicity; epidermal growth factor receptor;

XX tumour necrosis factor; lymphocyte; tetra valent antibody; cytostatic;

XX antiinflammatory; antipsoriasis; dermatological; antiulcer;

XX antiathermatic; antiarteriosclerotic; antineumatic; antibacterial;

XX antiarthritic; neuroprotective; immunosuppressive; antianemic;

XX antiallergic; antidiabetic; gene therapy; human.

XX Homo sapiens.

XX WO200177342-A1.

XX 18-OCT-2001.

XX 20-MAR-2001; 2001WO-US08928.

XX 11-APR-2000; 2000US-195819P.

PA (GETH) GENENTECH INC.

XX Miller KL, Presta LG;

PI WPI; 2002-049149/06.

DR Novel engineered antibody useful in therapeutic applications, contains

XX a dimerisation domain and three or more antigen binding sites -

XX Claim 11; Page 14; 186pp; English.

XX This invention relates to an isolated antibody comprising a

XX dimerisation domain and three or more antigen binding sites

XX amino-terminal to the domain. It is cytostatic, antiinflammatory,

XX antibacterial, immunosuppressive, antiallergic, an apoptosis

XX inducer, a vaccine and used in gene therapy. Along with a cytotoxic

XX agent, is useful for treating a disorder e.g. cancer in a mammal, for

XX inducing apoptosis of a cancer cell, and for killing a B cell or a cell

XX which overexpresses or expresses an ErbB receptor and for treating

XX benign and malignant tumours, inflammatory, angiogenic and immunological

XX disorders, autoimmune diseases, central nervous system inflammatory

XX disorders. The antibody is also useful for immunodiagnosis of various

XX diseases including cancer, for human therapy in redirected cytotoxicity,

XX and also useful as fibrinolytic agents or vaccine adjuvants, useful as

XX affinity purification agent, in diagnostic assays for detecting the

XX expression of antigen of interest in specific cells, tissue or

XX serum, and useful for blocking an immune response to a foreign antigen.

XX The antigen is internalised faster than a bivalent antibody by a

XX cell expressing an antigen to which the antibodies bind. The

XX antibody comprises three or four heavy chain variable domains which are

XX able to combine with three or four light chain variable domain

XX polypeptides to form three or four antigen binding sites directed

XX against the same antigen. This sequence represents the peptide

XX sequence of a Fd flexible linker which can be found on a polypeptide

XX chain of the antibodies.

SQ Sequence 4 AA;

Query Match 100.0%; Score 20; DB 23; Length 4;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSGS 4
|||

Db 1 GSGS 4

RESULT 10

ABG73964

XX ID ABG73964 standard; Peptide; 4 AA.

XX AC ABG73964;

XX DT 28-MAR-2003 (first entry)

XX DE Synthetic peptide linker sequence.

XX Hepatitis C; viral protease; anti-viral; tumour; linker;

XX virus; infection; antitumour; toxophore; human immunodeficiency virus;

XX HIV infection; herpes simplex virus; HSV; rhinovirus; NS3 protease.

XX Synthetic.

XX WO200287500-A2.

XX 07-NOV-2002.

XX 26-APR-2002; 2002WO-US13223.

XX 27-APR-2001; 2001US-286893P.

XX (NEWB-) NEWBIOTICS INC.

PI Cathers BF, Neuteboom STC, Shepard HM;
 XX WPI; 2003-167102/16.
 XX Novel synthetic viral prototoxophore for treating viral infections, has
 PT toxin moiety incorporated into substrate domain specific for viral
 PT enzyme, bound and modified by viral enzyme to get converted into
 PT toxophore -
 XX Example 1; Page 38; 66pp; English.
 XX This invention relates to a novel synthetic viral prototoxophore
 CC comprising a toxin moiety operatively incorporated into a substrate
 CC domain specific for a viral enzyme. This prototoxophore may be bound
 CC and modified by the viral enzyme thus converting it to a toxophore.
 CC Also disclosed in the invention is a method for enhancing the anti-viral
 CC effect of an antiviral agent, this method comprises contacting a cell,
 CC infected with a virus or is susceptible to infection, with a
 CC prototoxophore. The invention further comprises an assay to identify
 CC anti-viral agents, comprising contacting an infected cell with a
 CC candidate agent and comparing the ability of the agent to inhibit the
 CC growth or infectivity of the virus in the cell. The prototoxophores
 CC of the invention may have virucide or antitumour activity. The
 CC prototoxophores of the invention may be useful for reducing or
 CC inhibiting viral infectivity, by contacting a cell (e.g. lymphocyte,
 CC nerve cell, connective tissue cell, muscle cell or hepatocyte) which is
 CC infected with a virus or is susceptible to infection with a virus, with
 CC an effective amount of the prototoxophore. The cells are cell lines
 CC adapted to long term continuous culture or isolated from a subject.
 CC The prototoxophore is also useful for ameliorating the severity of a
 CC viral infection in a subject, where the virus is selected from human
 CC immunodeficiency virus (HIV), herpes simplex virus (HSV), rhinovirus and
 CC hepatitis virus, by administering an effective amount of the
 CC prototoxophore to the subject. The prototoxophores of the invention are
 CC also useful for treating tumours. The present sequence represents a
 CC synthetic linker peptide sequence used to create an NS3/NS4 fusion
 CC protein for use in the examples of the invention.
 XX SQ Sequence 4 AA;
 Query Match 100.0%; Score 20; DB 24; Length 4;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GSGS 4
 Db ||||
 1 GSGS 4
 RESULT 11
 AAW56059
 ID AAW56059 standard; peptide; 5 AA.
 XX AAW56059;
 AC AAW56059;
 XX 29-JUL-1998 (first entry)
 DT
 XX Homo sapiens.
 DE Mutant CD loop of the Ad5 fiber knob.
 XX Chimeric; adenovirus; fiber protein; binding; targeting; coat protein;
 KW constrained peptide motif; gene therapy; cancer; heart disease;
 KW autoimmune disorder.
 XX Synthetic.
 OS Mastadenovirus.
 XX WO9807865-A1.
 XX 26-FEB-1998.
 XX 21-AUG-1997; 97WO-US14719.
 XX 21-AUG-1996; 96US-0701124.

XX (GENV-) GENVEC INC.
 XX Kovesdi I, Roelvink PW, Wickham TJ;
 XX WPI; 1998-169169/15.
 XX Chimeric adenovirus fibre proteins - containing non-native amino
 PT acid sequence to provide for binding and entry into cells,
 PT especially for gene therapy
 XX Example 1; Page 45; 124pp; English.
 XX The present sequence represents a mutant CD loop of the Ad5 fiber
 CC knob which is used in an example from the present invention. The
 CC present invention describes a chimeric adenovirus fibre protein (AFP)
 CC containing a constrained non-native amino acid sequence. The non-native
 CC amino acid sequence allows the chimeric fibre (or a vector comprising
 CC the chimeric fibre) to more efficiently bind to and enter cells. The
 CC products can be used for gene therapy, for treating cancer, e.g.
 CC melanoma, glioma and lung cancers as well as genetic disorders, e.g.
 CC cystic fibrosis, haemophilia and muscular dystrophy as well as
 CC pathogenic infections, e.g. HIV, tuberculosis and hepatitis and also for
 CC heart disease, to e.g. prevent restenosis following angioplasty or to
 CC promote angiogenesis to reperfuse necrotic tissue, and in autoimmune
 CC disorders, e.g. Crohn's disease, colitis, rheumatoid arthritis, and
 CC Alzheimer's disease.
 XX SQ Sequence 5 AA;
 Query Match 100.0%; Score 20; DB 19; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GSGS 4
 Db ||||
 1 GSGS 4
 RESULT 12
 AAB07386
 ID AAB07386 standard; Protein; 5 AA.
 XX AAB07386;
 AC AAB07386;
 XX 17-JAN-2001 (first entry)
 DT
 XX Human alpha_d metal-binding region consensus sequence.
 DE Human; macrophage infiltration inhibition; alpha_d integrin;
 KW leukocyte integrin; leu-CAM; leukointegrin; immune response;
 KW inflammation; leukocyte adhesion deficiency; LAD; Type I diabetes;
 KW atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;
 KW lung inflammation; acute respiratory distress syndrome; Crohn's disease;
 KW rheumatoid arthritis; central nervous system injury; metal-binding.
 XX Homo sapiens.
 OS WO200029446-A1.
 XX 25-MAY-2000.
 XX 16-NOV-1999; 99WO-US27139.
 XX 16-NOV-1998; 98US-0193043.
 XX 08-JUL-1999; 99US-0350259.
 XX (ICOS-) ICOS CORP.
 XX Gallatin MW, Van Der Vlieten M;
 XX WPI; 2000-387751/33.

PT Use of novel anti-alpha integrin d monoclonal antibodies to inhibit
PT macrophage infiltration and reduce inflammation at central nervous
PT system injury sites -
XX
XX Example 5; Page 32; 270pp; English.
XX
XX Integrins are a class of membrane-associated molecules that participate
CC in cellular adhesion. Integrins are made up of an alpha subunit and a
CC beta subunit. One class of human integrins are restricted to expression
CC in white blood cells and have a common beta2 subunit; the leukocyte
CC integrins, leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins
CC have an important role in immune and inflammatory responses. Alpha d is a
CC beta2 integrin alpha subunit. The present sequence is the human alpha d
CC metal-binding region consensus sequence. This sequence is required for
CC ligand interaction. Alpha d gene/ protein may be useful in therapy for
CC diseases linked to aberrant alpha d function e.g. Type I diabetes,
CC atherosclerosis, multiple sclerosis, asthma, psoriasis, lung
CC inflammation, acute respiratory distress syndrome, rheumatoid arthritis
CC and leukocyte adhesion deficiency (LAD). In addition, anti-alpha d
CC monoclonal antibodies may be used in the inhibition of macrophage
CC infiltration at the site of a central nervous system injury. The
CC monoclonal antibodies can also be used to detect and diagnose Crohn's
CC disease.
XX
XX Sequence 5 AA;
XX
XX Query Match 100.0%; Score 20; DB 21; Length 5;
XX Best Local Similarity 100.0%; Pred. No. 9.3e+05;
XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 GSGS 4
XX ||||
XX Db 2 GSGS 5
XX
XX RESULT 13
XX ABJ05341
XX ID ABJ05341 standard; Peptide; 5 AA.
XX AC ABJ05341;
XX DT 08-NOV-2002 (first entry)
XX DE Target fusion peptide production method-related linker peptide.
XX
XX Target peptide production; fusion peptide; protease-susceptible linker;
XX parathyroid hormone; PTH; high expression rate;
XX pharmaceutical application.
XX Unidentified.
XX WO200236762-A1.
XX
XX 10-MAY-2002.
XX
XX 29-OCT-2001; 2001WO-JP09476.
XX
XX 30-OCT-2000; 2000JP-0331170.
XX
XX 27-JUN-2001; 2001JP-0195522.
XX
XX (TAKE) TAKEDA CHEM IND LTD.
XX
XX Yamada T, Suenaga M;
XX WPI; 2002-417275/44.
XX N-PSDB; ABT06839.
XX
XX Production of target peptide comprises cleavage of fusion peptide with
PT parathyroid hormone peptide for efficient manufacture of target peptide
PT without the need to remove N-terminal methionine -
XX
XX Disclosure; Page 16; 103pp; Japanese.

CC The invention comprises a method of producing a target peptide. The
CC C-terminal end of the target peptide is fused via a protease-susceptible
CC linker to parathyroid hormone (PTH) residues 1-34. The method of the
CC invention is useful for the clean and efficient production of a target
CC peptide at a high expression rate on an industrial scale without the need
CC to remove the N-terminal methionine from the product. The peptides
CC produced by the method of the invention are suitable for pharmaceutical
CC and other uses. The present peptide sequence was used in the invention.
XX
XX Sequence 5 AA;
XX
XX Query Match 100.0%; Score 20; DB 23; Length 5;
XX Best Local Similarity 100.0%; Pred. No. 9.3e+05;
XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 GSGS 4
XX ||||
XX Db 1 GSGS 4
XX
XX RESULT 14
XX AAO26595
XX ID AAO26595 standard; Peptide; 5 AA.
XX AC AAO26595;
XX DT 28-MAR-2003 (first entry)
XX DE Fusion protein related peptide #3.
XX KW Fungicide; fusion protein; cell surface layer; target cell; specificity;
XX fungal infection.
XX OS Unidentified.
XX PN JP2002253245-A.
XX PD 10-SEP-2002.
XX PF 28-FEB-2001; 2001JP-0055200.
XX PR 28-FEB-2001; 2001JP-0055200.
XX PA (TOYOW) TOYOTA CHUO KENKYUSHO KK.
XX DR WPI; 2003-150965/15.
XX PT A fusion protein useful for protecting an organism against fungi
XX comprises a first peptide site acting on cell surface of a target cell
XX and a second peptide site which interacts with cell surface -
XX Disclosure; Page 5; 15pp; Japanese.
XX
XX The invention relates to a novel fusion protein containing a first
XX peptide site which can act on the cell surface layer of a target cell and
XX a second peptide site which can combine to the constituent of the cell
XX surface layer of the target cell. The fusion protein is useful for
XX protecting an organism against fungal infection. The method can be used
XX for enhancing specificity of a membrane-acting peptide to a specific
XX target cell. This sequence represents a peptide relating to the fusion
XX protein of the invention.
XX
XX Sequence 5 AA;
XX
XX Query Match 100.0%; Score 20; DB 24; Length 5;
XX Best Local Similarity 100.0%; Pred. No. 9.3e+05;
XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 GSGS 4
XX ||||
XX Db 1 GSGS 4

RESULT 15

AAW56062
ID AAW56062 standard; peptide; 6 AA.XX AC
XX AAW56062;

XX DT 29-JUL-1998 (first entry)

XX DE Mutant FG loop of the Ad5 fiber knob.

XX KW Chimeric; adenovirus; fiber protein; binding; targeting; coat protein;
XX KW constrained peptide motif; gene therapy; cancer; heart disease;
XX KW autoimmune disorder.

XX OS Synthetic.

XX OS Mastadenovirus.

XX FN WO9807865-A1.

XX PD 26-FEB-1998.

XX PF 21-AUG-1997; 97WO-US14719.

XX PR 21-AUG-1996; 96US-0701124.

XX PA (GENV-) GENVEC INC.

XX PI Kovesdi I, Roelvink PW, Wickham TJ;

XX DR WPI; 1998-169169/15.

XX PT Chimeric adenovirus fibre proteins - containing non-native amino
XX PT acid sequence to provide for binding and entry into cells,
XX PT especially for gene therapy

XX BS Example 1; Page 46; 124pp; English.

XX CC The present sequence represents a mutant FG loop of the Ad5 fiber
 CC knob, which is used in an example from the present invention. The
 CC present invention describes a chimeric adenovirus fibre protein (AFP)
 CC containing a constrained non-native amino acid sequence. The non-native
 CC amino acid sequence allows the chimeric fibre (or a vector comprising
 CC the chimeric fibre) to more efficiently bind to and enter cells. The
 CC products can be used for gene therapy, for treating cancer, e.g.
 CC melanoma, glioma and lung cancers as well as genetic disorders, e.g.
 CC cystic fibrosis, haemophilia and muscular dystrophy as well as
 CC pathogenic infections, e.g. HIV, tuberculosis and hepatitis and also for
 CC heart disease, to e.g. prevent restenosis following angioplasty or to
 CC promote angiogenesis to reperfuse necrotic tissue, and in autoimmune
 CC disorders, e.g. Crohn's disease, colitis, rheumatoid arthritis, and
 CC Alzheimer's disease.

XX SQ Sequence 6 AA;

Query Match 100.0%; Score 20; DB 19; Length 6;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSGS 4

|||||
 Db 1 GSGS 4

Search completed: February 11, 2004, 22:17:49
 Job time : 79 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 21:58:40 ; Search time 38 Seconds
(without alignments)
10.123 Million cell updates/sec

Title: US-09-813-341-10

Perfect score: 20
Sequence: 1 GSGS 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76:**

1: Pir1:*
2: Pir2:*
3: Pir3:*
4: Pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	6	PT0280	Ig heavy chain CRD
2	20	100.0	17	PH1367	Ig heavy chain DJ
3	20	100.0	17	H53284	T-cell receptor be
4	20	100.0	18	PH1323	Ig heavy chain DJ
5	20	100.0	19	PT0332	Ig heavy chain CRD
6	20	100.0	21	PT0227	Ig heavy chain CRD
7	20	100.0	23	S70339	napin small chain
8	20	100.0	23	A32473	histidine-rich pro
9	20	100.0	31	T27844	hypothetical prote
10	20	100.0	32	D87579	hypothetical prote
11	20	100.0	34	H30607	Ig kappa chain V-I
12	20	100.0	39	G32529	Ig lambda chain V
13	20	100.0	42	A31918	cathepsin D (EC 3.
14	20	100.0	42	PQ0457	coat protein - zuc
15	20	100.0	42	PQ0458	coat protein - zuc
16	20	100.0	42	PQ0456	coat protein vsaH [
17	20	100.0	42	A99578	lipoprotein
18	20	100.0	43	T02348	hypothetical prote
19	20	100.0	44	C47193	1-caldesmon I (Hel
20	20	100.0	46	A45758	hypothetical glucoc
21	20	100.0	51	I57670	folliotropin recept
22	20	100.0	53	S12520	core protein A1 -
23	20	100.0	54	S34093	Ig kappa chain V r
24	20	100.0	54	S40381	Ig kappa chain V-I
25	20	100.0	54	A25521	Ig kappa chain V r
26	20	100.0	54	JT0521	Ig kappa chain V-I
27	20	100.0	54	AB2291	hypothetical prote
28	20	100.0	56	S45372	peptidylprolyl iso
29	20	100.0	57	S09493	Ig heavy chain pre

30 20 100.0 58 2 C82527
31 20 100.0 60 2 B40128
32 20 100.0 64 2 PC1237
33 20 100.0 64 2 B41286
34 20 100.0 64 2 A86333
35 20 100.0 64 2 C38601
36 20 100.0 65 2 B38601
37 20 100.0 67 2 PH1091
38 20 100.0 67 2 S24216
39 20 100.0 67 2 S20172
40 20 100.0 67 2 S20172
41 20 100.0 68 2 S26474
42 20 100.0 68 2 JQ2005
43 20 100.0 69 1 MIEC77
44 20 100.0 69 2 PH1080
45 20 100.0 69 2 A64961

hypothetical prote
probable antigen 1
peptidylprolyl iso
granulocyte-macrop
hypothetical prote
Ig kappa chain V r
Ig kappa chain V r
Ig light chain V r
Ig kappa chain - m
hypothetical prote
hypothetical prote
Ig kappa chain V r
hypothetical 7.5K
microcin B17 precu
Ig light chain V r
outer membrane por

ALIGNMENTS

RESULT 1

PT0280

Ig heavy chain CRD3 region (clone 4-91B) - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996

C:Accession: PT0280

Riyamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.

J. Exp. Med. 173, 395-407, 1991

A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity and j.

A:Reference number: PT0222; MUID:91108337; PMID:1899102

A:Accession: PT0280

A:Molecule type: DNA

A:Residues: 1-6 <YAM>

A:Experimental source: B lymphocyte

C:Keywords: heterotetramer; immunoglobulin

Query Match 100.0%; Score 20; DB 2; Length 6;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSGS 4

DB 1 GSGS 4

RESULT 2

PH1367

Ig heavy chain DJ region (clone C111-106B) - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999

C:Accession: PH1367

R.Wasserman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.

J. Exp. Med. 176, 1577-1581, 1992

A:Title: Predominance of fetal type DJH joining in young children with B precursor lymph

A:Reference number: PH1302; MUID:93094761; PMID:1460419

A:Accession: PH1367

A:Molecule type: DNA

A:Residues: 1-17 <WAS>

C:Keywords: heterotetramer; immunoglobulin

Query Match 100.0%; Score 20; DB 2; Length 17;

Best Local Similarity 100.0%; Pred. No. 1.6e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSGS 4

DB 3 GSGS 6

RESULT 3

H53284

T-cell receptor beta 2 chain J region, Jbeta2.6 - rabbit

C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 02-May-1994 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
C:Accession: H53284
R:Harindranath, N.; Alexander, C.B.; Mage, R.G.
Mol. Immunol. 28, 881-888, 1991
A:Title: Evolutionarily conserved organization and sequences of germline diversity and
A:Reference number: A53284; MUID:91342695; PMID:1678859
A:Accession: H53284
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-17 <HAR>
A:Cross-references: GB:S60737; NID:G233916; PID:AB19524.1; PID:G233924
A:Note: sequence extracted from NCBI backbone (NCBI:60737, NCBI:60746)
C:Keywords: T-cell receptor

Query Match 100.0%; Score 20; DB 2; Length 17;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GSGS 4
Db 9 GSGS 12

RESULT 4

PH1323
Ig heavy chain DJ region (clone C174-113) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PH1323
R:Wasserman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.
J. Exp. Med. 176, 1577-1581, 1992
A:Title: Predominance of fetal type DJH joining in young children with B precursor lymphoma
A:Reference number: PH1302; MUID:93094761; PMID:1460419
A:Accession: PH1323
A:Molecule type: DNA
A:Residues: 1-18 <WAS>
A:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 100.0%; Score 20; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GSGS 4
Db 4 GSGS 7

RESULT 5

PT0332
Ig heavy chain CRD3 region (clone J2-139) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C:Accession: PT0332
R:Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity and
A:Reference number: PT0222; MUID:91108337; PMID:1899102
A:Accession: PT0332
A:Molecule type: DNA
A:Residues: 1-19 <YAM>
A:Experimental source: B lymphocyte
C:Keywords: heterotetramer; immunoglobulin

Query Match 100.0%; Score 20; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GSGS 4
Db 10 GSGS 13

RESULT 6

PT0227
Ig heavy chain CDR3 region (clone 1-106) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C:Accession: PT0227
R:Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity and
A:Reference number: PT0222; MUID:91108337; PMID:1899102
A:Accession: PT0227
A:Molecule type: DNA
A:Residues: 1-21 <YAM>
A:Experimental source: B lymphocyte
C:Keywords: heterotetramer; immunoglobulin

Query Match 100.0%; Score 20; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GSGS 4
Db 7 GSGS 10

RESULT 7

S70339
napin small chain S4 - Swedish turnip (fragments)
C:Species: Brassica napus var. rapifera (Swedish turnip, rutabaga)
C:Date: 19-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 07-May-1999
C:Accession: S70339
R:Neumann, G.M.; Condron, R.; Thomas, I.; Polya, G.M.
Biochim. Biophys. Acta 1295, 23-33, 1996
A:Title: Purification and sequencing of multiple forms of Brassica napus seed napin small
A:Reference number: S70336; MUID:96283790; PMID:8679670
A:Accession: S70339
A>Status: Preliminary
A:Molecule type: protein
A:Residues: 1-3; 4-11; 12-23 <NEU>
C:Superfamily: wheat alpha-amylase inhibitor

Query Match 100.0%; Score 20; DB 2; Length 23;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GSGS 4
Db 17 GSGS 20

RESULT 8

A32473
histidine-rich protein C - liver fluke (fragment)
C:Species: Fasciola hepatica (liver fluke)
C:Date: 25-Sep-1989 #sequence_revision 12-May-1994 #text_change 15-Oct-1999
C:Accession: A32473
R:Waite, J.H.; Rice-Ficht, A.C.
Biochemistry 28, 6104-6110, 1989
A:Title: A histidine-rich protein from the vitellaria of the liver fluke Fasciola hepatica
A:Reference number: A32473; MUID:89375343; PMID:2775756
A:Accession: A32473
A:Molecule type: protein
A:Residues: 1-23 <WAI>
A:Note: 22-Gly, 22-Ser, 23-Gly, and 23-Ser were also found
C:Keywords: egg yolk
F;6,8,12,16/Modified site: 3', 4'-dihydroxyphenylalanine (Tyr) #status experimental

Query Match 100.0%; Score 20; DB 2; Length 23;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GSGS 4

Db 17 GSGS 20
|||||
RESULT 9
T27844
hypothetical protein ZK402.4 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T27844
R:Favella, T.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of *C. elegans* cosmid ZK402.
A:Reference number: Z20429
A:Accession: T27844
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-31 <FAV>
A:Cross-references: EMBL:U40956; PIDN:AAA81755.1; CESP:ZK402.4
C:Genetics:
A:Gene: CESP:ZK402.4
Query Match 100.0%; Score 20; DB 2; Length 31;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GSGS 4
|||||
Db 24 GSGS 27
RESULT 10
D87579
hypothetical protein CC2665 [imported] - *Caulobacter crescentus*
C:Species: *Caulobacter crescentus*
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: D87579
R:Nierman, W.C.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
B.; Laub, M.T.; Tabor, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of *Caulobacter crescentus*.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: D87579
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-32 <STO>
A:Cross-references: GB:AE005673; NID:gl3424248; PIDN:AAK24632.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC2665
Query Match 100.0%; Score 20; DB 2; Length 32;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GSGS 4
|||||
Db 25 GSGS 28
RESULT 11
H30607
Ig kappa chain V-III region (Bla) - human (fragment)
C:Species: *Homo sapiens* (man)
C:Date: 29-Jun-1989 #sequence_revision 29-Jun-1989 #text_change 30-May-1997
C:Accession: H30607
R:Goni, F.R.; Chen, P.P.; McGinnis, D.; Arjonilla, M.L.; Fernandez, J.; Carson, D.; Sold
J. Immunol. 142, 3158-3163, 1989
A:Title: Structural and idiotypic characterization of the L chains of human IGM autoanti
A:Reference number: A30601; MUID:89215279; PMID:2496160
A:Accession: H30607
A:Status: preliminary
A:Molecule type: protein

A:Residues: 1-34 <GON>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
Query Match 100.0%; Score 20; DB 2; Length 34;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GSGS 4
|||||
Db 18 GSGS 21
RESULT 12
Ig lambda chain V region (clone pH2) - rabbit (fragment)
C:Species: *Oryctolagus cuniculus* (domestic rabbit)
C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 30-May-1997
C:Accession: G32529
R:Hayzer, D.J.; Duvoisin, R.M.; Jaton, J.C.
Biochem. J. 245, 691-697, 1987
A:Title: cDNA clones encoding rabbit immunoglobulin lambda chains. Evidence for length v
A:Reference number: A90338; MUID:88024122; PMID:3117050
A:Accession: G32529
A:Molecule type: mRNA
A:Residues: 1-39 <HAY>
A:Cross-references: GB:M25620
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
Query Match 100.0%; Score 20; DB 2; Length 39;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GSGS 4
|||||
Db 20 GSGS 23
RESULT 13
A31918
cathepsin D (BC 3.4.23.5) - bovine (fragment)
C:Species: *Bos primigenius taurus* (cattle)
C:Date: 21-May-1990 #sequence_revision 31-Dec-1991 #text_change 01-Nov-1996
C:Accession: A31918
R:Yonezawa, S.; Takahashi, T.; Wang, X.; Wong, R.N.S.; Hartsuck, J.A.; Tang, J.
J. Biol. Chem. 263, 16504-16511, 1988
A:Title: Structures at the proteolytic processing region of cathepsin D.
A:Reference number: A92681; MUID:89034127; PMID:3182800
A:Accession: A31918
A:Molecule type: protein
A:Residues: 1-42 <YON>
C:Superfamily: pepsin
C:Keywords: aspartic proteinase; glycoprotein; hydrolase; lysosome
F:1-30/Product: cathepsin D light chain (fragment) #status experimental <LCH>
F:18-42/Product: cathepsin D, single-chain form (fragment) #status experimental <MAT>
F:33-42/Product: cathepsin D heavy chain (fragment) #status experimental <HCH>
F:1.28/Binding site: carbonyl site: Asn (covalent) #status predicted
Query Match 100.0%; Score 20; DB 2; Length 42;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GSGS 4
|||||
Db 10 GSGS 13
RESULT 14
PQ0457
coat protein - zucchini yellow mosaic virus (strain NATmv) (fragment)
C:Species: zucchini yellow mosaic virus, ZYMV
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Nov-2000

C;Accession: PQ0457
R;Gal-On, A.; Antignus, Y.; Rosner, A.; Raccach, B.
J. Gen. Virol. 73, 2183-2187, 1992
A;Title: A zucchini yellow mosaic virus coat protein gene mutation restores aphid transmissibility
A;Reference number: PQ0456; MUID:93019038; PMID:1402810
A;Accession: PQ0457
A;Molecule type: mRNA
A;Residues: 1-42 <GAL>
A;Experimental source: strain ZYMV-NATmv
C;Superfamily: tobacco etch virus genome polyprotein
C;Keywords: coat protein

Query Match 100.0%; Score 20; DB 2; Length 42;
Best Local Similarity 100.0%; Pred.No. 4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSGS 4
|||
Db 29 GSGS 32

RESULT 15
PQ0458
coat protein - zucchini yellow mosaic virus (strain AT) (fragment)
C;Species: zucchini yellow mosaic virus, ZYMV
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Nov-2000
C;Accession: PQ0458
R;Gal-On, A.; Antignus, Y.; Rosner, A.; Raccach, B.
J. Gen. Virol. 73, 2183-2187, 1992
A;Title: A zucchini yellow mosaic virus coat protein gene mutation restores aphid transmissibility
A;Reference number: PQ0456; MUID:93019038; PMID:1402810
A;Accession: PQ0458
A;Molecule type: mRNA
A;Residues: 1-42 <GAL>
A;Cross-references: GB:S46009; NID:g257075; PIDN:AB23550.1; PID:g257076
A;Experimental source: strain ZYMV-AT
C;Superfamily: tobacco etch virus genome polyprotein
C;Keywords: coat protein

Query Match 100.0%; Score 20; DB 2; Length 42;
Best Local Similarity 100.0%; Pred.No. 4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSGS 4
|||
Db 29 GSGS 32

Search completed: February 11, 2004, 22:21:08
Job time : 45 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 19:08:34 ; Search time 24 Seconds
(without alignments)
7.838 Million cell updates/sec

Title: US-09-813-341-10
Perfect score: 20
Sequence: 1 GSGS 4

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	20	100.0	63	1 MIP_BOTAS	P81077 bothrops as
2	20	100.0	63	1 PER_DROIM	Q03294 drosophila
3	20	100.0	63	1 PER_DROMS	Q04535 drosophila
4	20	100.0	64	1 CYPM BOVIN	P30404 bos taurus
5	20	100.0	64	1 MTCU HELPO	P55947 helix pomat
6	20	100.0	65	1 PER_DROMO	Q03295 drosophila
7	20	100.0	66	1 PER_DROSA	Q04536 drosophila
8	20	100.0	69	1 MCBF ECOLI	P05834 escherichia
9	20	100.0	75	1 MCHB ECOLI	Q9RM53 escherichia
10	20	100.0	80	1 KORB STRLI	P22404 streptomyce
11	20	100.0	83	1 V187 BPT3	P10302 bacterioph
12	20	100.0	84	1 CRH_EACHD	Q9K708 bacillus ha
13	20	100.0	85	1 IATP YEAST	P01097 saccharomyc
14	20	100.0	86	1 PER_DRORO	Q03296 drosophila
15	20	100.0	86	1 RK27 GUITH	Q78430 guillardia
16	20	100.0	86	1 RK27 PORPU	P51210 porphyra pu
17	20	100.0	87	1 IM13 ARATH	Q9KH48 arabidopsis
18	20	100.0	87	1 RL27 TREPA	O83725 treponema p
19	20	100.0	88	1 IATP CABEL	P37209 caenorhabdi
20	20	100.0	88	1 PER_DROTE	Q26287 drosophila
21	20	100.0	91	1 RK15 SPIOL	P22798 spinacia ol
22	20	100.0	92	1 KV09 RABIT	P01690 cryptotolagus
23	20	100.0	93	1 R31B PSESM	Q87XJ5 pseudomonas
24	20	100.0	94	1 ESA6 MYCTU	Q57165 mycobacteri
25	20	100.0	94	1 KV11 RABIT	P01692 cryptotolagus
26	20	100.0	95	1 CXA3 CERLA	P01527 cerebratulu
27	20	100.0	95	1 IL3B HUMAN	Q9Y514 homo sapien
28	20	100.0	96	1 HOL1_BPA18	Q37975 bacterioph
29	20	100.0	96	1 HOL1_BPA50	Q37977 bacterioph
30	20	100.0	97	1 KCH2 FIG	Q9TU14 sus scrofa
31	20	100.0	98	1 2C DICDI	P15648 dictyosteli
32	20	100.0	100	1 RK27 CYACA	O19885 cyanidium c
33	20	100.0	101	1 CH10_HUMAN	Q04984 homo sapien

34 20 100.0 101 1 VATF METMA Q60185 methanosarc
35 20 100.0 103 1 CXKL_BUNNM Q9PW19 bungarus mu
36 20 100.0 103 1 MCEA_KLEPN Q9Z4N4 klebsiella
37 20 100.0 104 1 KACE6_RABIT P03984 cryptotolagus
38 20 100.0 105 1 PLA2_LEIBR Q44010 leishmania
39 20 100.0 105 1 PLA2_LEIDO Q43940 leishmania
40 20 100.0 106 1 RLA2_LELIN Q06383 leishmania
41 20 100.0 106 1 Y116_ADE02 P01685 cryptotolagus
42 20 100.0 107 1 KV04_RABIT P01596 homo sapien
43 20 100.0 107 1 KVID_HUMAN P01675 mus musculu
44 20 100.0 107 1 KV6A_MOUSE P01676 mus musculu
45 20 100.0 107 1 KV6B_MOUSE

ALIGNMENTS

RESULT 1

MIP_BOTAS
ID MIP_BOTAS STANDARD; PRT; 63 AA.
AC P81077;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Myotoxin inhibitor protein MIP.
OS Bothrops asper (TerCiopelel).
OC Eukaryota, Metazoa, Chordata, Craniata, Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Bothrops.
OX NCBI_TaxID=8722;
RN [1]
RP SEQUENCE.
RC TISSUE=Blood;
RX MEDLINE=97439729; PubMed=9307037;
RA Liiano S., Lomonte B., Fox J.W., Gutierrez J.M.;
RT "Biochemical characterization and pharmacological properties of a
RT phospholipase A2 myotoxin inhibitor from the plasma of the snake
RT Bothrops asper.";
RL Biochem. J. 326:853-859(1997).
CC -!- FUNCTION: BINDS TO AND NEUTRALIZES THE ACTIVITIES OF BASIC
CC -!- PHOSPHOLIPASE A2 (PLA2) MYOTOXIN ISOFORMS.
CC -!- SUBUNIT: OLIGOMER COMPOSED OF FIVE 23-25 KDa SUBUNITS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
KW Glycoprotein; Plasma; Lectin.
SQ SEQUENCE 63 AA; 7154 MW; F1B6AF75D8489CE6 CRC64;

Query Match Similarity 100.0%; Score 20; DB 1; Length 63;
Best Local Similarity 100.0%; Pred. No. 2.8e+02; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0;

QY 1 GSGS 4
DB 41 GSGS 44

RESULT 2

PER_DROIM
ID PER_DROIM STANDARD; PRT; 63 AA.
AC Q03294;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-JUL-1999 (Rel. 36, Last annotation update)
DE Period circadian protein (fragment).
GN PER.
OS Drosophila immigrans (Fruit fly).
OC Eukaryota, Metazoa, Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7250;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93196482; PubMed=8450754;

RA Peixoto A.A., Campesan S., Costa R.H., Kyriacou C.P.;
 RT "Molecular evolution of a repetitive region within the per gene of
 RL Drosophila.";
 CC Mol. Biol. Evol. 10:127-139 (1993).
 CC -!- FUNCTION: ESSENTIAL FOR BIOLOGICAL CLOCK FUNCTIONS. DETERMINES THE
 CC PERIOD LENGTH OF CIRCADIAN AND ULTRADIAN RHYTHMS; AN INCREASE IN
 CC PER DOSAGE LEADS TO SHORTENED CIRCADIAN RHYTHMS AND A DECREASE
 CC LEADS TO LENGTHENED CIRCADIAN RHYTHMS. ESSENTIAL FOR THE CIRCADIAN
 CC RHYTHMIC COMPONENT OF THE MALE COURTSHIP SONG THAT ORIGINATES IN
 CC THE THORACIC NERVOUS SYSTEM. THE BIOLOGICAL CYCLE DEPENDS ON THE
 CC RHYTHMIC COMPONENT OF THE MALE COURTSHIP SONG THAT ORIGINATES IN
 CC THE THORACIC NERVOUS SYSTEM. THE BIOLOGICAL CYCLE DEPENDS ON THE
 CC RHYTHMIC FORMATION AND NUCLEAR LOCALIZATION OF THE TIM-PER
 CC COMPLEX. LIGHT INDUCES THE DEGRADATION OF TIM, WHICH PROMOTES
 CC ELIMINATION OF PER. NUCLEAR ACTIVITY OF THE HETERODIMER
 CC COORDINATIVELY REGULATES PER AND TIM TRANSCRIPTION THROUGH A
 CC NEGATIVE FEEDBACK LOOP. BEHAVES AS A NEGATIVE ELEMENT IN CIRCADIAN
 CC TRANSCRIPTIONAL INHIBITION (BY SIMILARITY).
 CC -!- SUBUNIT: FORMS HETERODIMER WITH TIMELESS (TIM); THE COMPLEX THEN
 CC INDIRECT TRANSCRIPTIONAL INHIBITION (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: NUCLEAR AT SPECIFIC PERIODS OF THE DAY.
 CC FIRST ACCUMULATES IN THE PERINUCLEAR REGION ABOUT ONE HOUR BEFORE
 CC TRANSLLOCATION INTO THE NUCLEUS. INTERACTION WITH TIM IS REQUIRED
 CC FOR NUCLEAR LOCALIZATION (BY SIMILARITY).
 CC -!- PTM: PHOSPHORYLATED WITH A CIRCADIAN RHYTHMICITY. PROBABLY BY THE
 CC DOUBLE-TIME PROTEIN (DBT). PHOSPHORYLATION COULD BE IMPLICATED IN
 CC THE STABILITY OF PER MONOMER AND IN THE FORMATION OF HETERODIMER
 CC PER-TIM (BY SIMILARITY).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L06337; AAA28760.1; -;
 DR FlyBase; Fgn001242; Dmimlper.
 KW Biological rhythms; Repeat; Nuclear protein; Phosphorylation.
 FT NON_TER 1
 FT NON_TER 63 63
 FT NON_TER 63 63
 SQ SEQUENCE 63 AA; 5820 MW; CFB6063B92CA6AB9 CRC64;
 Query Match 100.0%; Score 20; DB 1; Length 63;
 Best Local Similarity 100.0%; Pred. No. 2.8e+02; Indels 0; Gaps 0;
 Matches 4; Conservative 0; Mismatches 0;
 QY 1 GSGS 4
 Db 5 GSGS 8
 RESULT 3
 PER DROMS STANDARD; PRT; 63 AA.
 ID _PER DROMS STANDARD; PRT; 63 AA.
 AC Q04535;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Period circadian protein (fragment).
 DE PER.
 GN Drosophila mediostriata (Fruit fly).
 OS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 CC Ephydroidea; Drosophilidae; Drosophila.
 CX NCBI_TaxID=7269;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=93196482; PubMed=8450754;
 RX Peixoto A.A., Campesan S., Costa R.H., Kyriacou C.P.;
 RA "Molecular evolution of a repetitive region within the per gene of
 RT Drosophila.";

RL Mol. Biol. Evol. 10:127-139 (1993).
 CC -!- FUNCTION: ESSENTIAL FOR BIOLOGICAL CLOCK FUNCTIONS. DETERMINES THE
 CC PERIOD LENGTH OF CIRCADIAN AND ULTRADIAN RHYTHMS; AN INCREASE IN
 CC PER DOSAGE LEADS TO SHORTENED CIRCADIAN RHYTHMS AND A DECREASE
 CC LEADS TO LENGTHENED CIRCADIAN RHYTHMS. ESSENTIAL FOR THE CIRCADIAN
 CC RHYTHMIC COMPONENT OF THE MALE COURTSHIP SONG THAT ORIGINATES IN
 CC THE THORACIC NERVOUS SYSTEM. THE BIOLOGICAL CYCLE DEPENDS ON THE
 CC RHYTHMIC FORMATION AND NUCLEAR LOCALIZATION OF THE TIM-PER
 CC COMPLEX. LIGHT INDUCES THE DEGRADATION OF TIM, WHICH PROMOTES
 CC ELIMINATION OF PER. NUCLEAR ACTIVITY OF THE HETERODIMER
 CC COORDINATIVELY REGULATES PER AND TIM TRANSCRIPTION THROUGH A
 CC NEGATIVE FEEDBACK LOOP. BEHAVES AS A NEGATIVE ELEMENT IN CIRCADIAN
 CC TRANSCRIPTIONAL INHIBITION (BY SIMILARITY).
 CC -!- SUBUNIT: FORMS HETERODIMER WITH TIMELESS (TIM); THE COMPLEX THEN
 CC INDIRECT TRANSCRIPTIONAL INHIBITION (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: NUCLEAR AT SPECIFIC PERIODS OF THE DAY.
 CC FIRST ACCUMULATES IN THE PERINUCLEAR REGION ABOUT ONE HOUR BEFORE
 CC TRANSLLOCATION INTO THE NUCLEUS. INTERACTION WITH TIM IS REQUIRED
 CC FOR NUCLEAR LOCALIZATION (BY SIMILARITY).
 CC -!- PTM: PHOSPHORYLATED WITH A CIRCADIAN RHYTHMICITY. PROBABLY BY THE
 CC DOUBLE-TIME PROTEIN (DBT). PHOSPHORYLATION COULD BE IMPLICATED IN
 CC THE STABILITY OF PER MONOMER AND IN THE FORMATION OF HETERODIMER
 CC PER-TIM (BY SIMILARITY).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L06338; AAA28761.1; -;
 DR FlyBase; Fgn0012531; Dmeslper.
 KW Biological rhythms; Repeat; Nuclear protein; Phosphorylation.
 FT NON_TER 1
 FT NON_TER 63 63
 FT NON_TER 63 63
 SQ SEQUENCE 63 AA; 5943 MW; D800662649BA6AB9 CRC64;
 Query Match 100.0%; Score 20; DB 1; Length 63;
 Best Local Similarity 100.0%; Pred. No. 2.8e+02; Indels 0; Gaps 0;
 Matches 4; Conservative 0; Mismatches 0;
 QY 1 GSGS 4
 Db 5 GSGS 8
 RESULT 4
 CYPM BOVIN STANDARD; PRT; 64 AA.
 ID _CYPM BOVIN STANDARD; PRT; 64 AA.
 AC P30404;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Peptidyl-prolyl cis-trans isomerase, mitochondrial (EC 5.2.1.8)
 DE (PPIase) (Rotamase) (Cyclophilin F) (Fragments).
 DE PPIF OR CYP3.
 GN Bos taurus (Bovine).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 CX NCBI_TaxID=9913;
 RN [1]
 RN SEQUENCE.
 RP TISSUE=Heart;
 RC MEDLINE=93176190; PubMed=7679902;
 RX Inoue T., Yoshida Y., Isaka Y., Tagawa K.;
 RA "Isolation of mitochondrial cyclophilin from bovine heart.";
 RT Biochem. Biophys. Res. Commun. 190:857-863 (1993).
 CC -!- FUNCTION: PPIases accelerate the folding of proteins. It catalyzes

the cis-trans isomerization of proline imidic peptide bonds in oligopeptides.
 -1- CATALYTIC ACTIVITY: Peptidylproline (omega=180) = peptidylproline (omega=0)
 -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
 PIR: PC1237; PC1237.
 DR InterPro: IPR002130; CSA_PPIase.
 DR PROSITE: PS00170; CSA_PPIase 1; PARTIAL.
 DR PROSITE: PS00072; CSA_PPIase 2; PARTIAL.
 KW Cyclosporin; Isomerase; Rotamase; Multigene family; Mitochondrion.
 FT NON_CONS 20 21
 FT NON_CONS 30 31
 FT NON_CONS 50 51
 FT NON_TER 64 64
 SQ SEQUENCE 64 AA; 6472 MW; 80926AF5625B9E5 CRC64;
 Query Match 100.0%; Score 20; DB 1; Length 64;
 Best Local Similarity 100.0%; Pred. No. 2.9e+02; Indels 0; Gaps 0;
 Matches 4; Conservative 0; Mismatches 0;
 QY 1 GSGS 4
 Db 4 GSGS 7
 RESULT 5
 MTCU HELPO STANDARD; PRT; 64 AA.
 ID P55947;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Copper-metallothionein (Cu-MT).
 OS Helix pomatia (Roman snail) (Edible snail).
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Stylommatophora;
 OC Sigurethra; Helicoidea; Helicidae; Helix.
 OX NCBI_TaxID=6536;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=97373947; PubMed=9230430;
 RA Dalling R., Berger B., Hunziker P.E., Kaegi J.H.R.;
 RT "Metallothionein in snail Cd and Cu metabolism."
 RL Nature 388:237-238(1997).
 CC -1- FUNCTION: THE METALLOTHIONEINS ARE INVOLVED IN THE CELLULAR SEQUESTRATION OF TOXIC METAL IONS AND REGULATION OF ESSENTIAL TRACE ELEMENTS. THIS ISOFORM BINDS EXCLUSIVELY COPPER.
 CC -1- DOMAIN: 14 CYSTEINE RESIDUES ARE ARRANGED IN C-X-C GROUPS. THESE ARE THOUGHT TO BE THE METAL-BINDING SITES IN OTHER METALLOTHIONEINS.
 CC -1- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 2.
 DR HSP; P05106; IJ02.
 DR InterPro: IPR002400; GFCysKNOT.
 DR PRINTS; PR00438; GFCYSKNOT.
 KW Metal-binding; Metal-thiolate cluster; Copper; Acetylation.
 FT MOD_RES 1 1
 FT METAL 7 7
 FT METAL 11 11
 FT METAL 16 16
 FT METAL 18 18
 FT METAL 22 22
 FT METAL 24 24
 FT METAL 28 28
 FT METAL 30 30
 FT METAL 33 33
 FT METAL 36 36
 FT METAL 38 38
 FT METAL 43 43
 FT METAL 45 45
 FT METAL 49 49
 FT METAL 55 55
 FT METAL 57 57

FT METAL 61 61 COPPER.
 FT METAL 63 63 COPPER.
 SQ SEQUENCE 64 AA; 6205 MW; 96CC1998B7E12297 CRC64;
 Query Match 100.0%; Score 20; DB 1; Length 64;
 Best Local Similarity 100.0%; Pred. No. 2.9e+02; Indels 0; Gaps 0;
 Matches 4; Conservative 0; Mismatches 0;
 QY 1 GSGS 4
 Db 51 GSGS 54
 RESULT 6
 PER_DROMO STANDARD; PRT; 65 AA.
 ID Q03295;
 AC Q03295;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Period circadian protein (Fragment).
 GN PER.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7230;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93196482; PubMed=8450754;
 RA Peixoto A.A., Campesan S., Costa R.H., Kyriacou C.P.;
 RT "Molecular evolution of a repetitive region within the per gene of Drosophila."
 RL Mol. Biol. Evol. 10:127-139(1993).
 CC -1- FUNCTION: ESSENTIAL FOR BIOLOGICAL CLOCK FUNCTIONS. DETERMINES THE PERIOD LENGTH OF CIRCADIAN AND ULTRADIAN RHYTHMS; AN INCREASE IN PER DOSAGE LEADS TO SHORTENED CIRCADIAN RHYTHMS AND A DECREASE LEADS TO LENGTHENED CIRCADIAN RHYTHMS. ESSENTIAL FOR THE CIRCADIAN RHYTHMICITY OF LOCOMOTOR ACTIVITY, ECLOSION BEHAVIOR, AND FOR THE RHYTHMIC COMPONENT OF THE MALE COURTSHIP SONG THAT ORIGINATES IN THE THORACIC NERVOUS SYSTEM. THE BIOLOGICAL CYCLE DEPENDS ON THE RHYTHMIC FORMATION AND NUCLEAR LOCALIZATION OF THE TIM-PER COMPLEX. LIGHT INDUCES THE DEGRADATION OF TIM, WHICH PROMOTES ELIMINATION OF PER. NUCLEAR ACTIVITY OF THE HETERODIMER COORDINATIVELY REGULATES PER AND TIM TRANSCRIPTION THROUGH A NEGATIVE FEEDBACK LOOP. BEHAVES AS A NEGATIVE ELEMENT IN CIRCADIAN TRANSCRIPTIONAL LOOP. DOES NOT APPEAR TO BIND DNA, SUGGESTING INDIRECT TRANSCRIPTIONAL INHIBITION (BY SIMILARITY).
 CC -1- SUBUNIT: FORMS HETERODIMER WITH TIMELESS (TIM); THE COMPLEX THEN TRANSLOCATES INTO THE NUCLEUS (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR AT SPECIFIC PERIODS OF THE DAY. FIRST ACCUMULATES IN THE PERINUCLEAR REGION ABOUT ONE HOUR BEFORE TRANSLOCATION INTO THE NUCLEUS. INTERACTION WITH TIM IS REQUIRED FOR NUCLEAR LOCALIZATION (BY SIMILARITY).
 CC -1- PTM: PHOSPHORYLATED WITH A CIRCADIAN RHYTHMICITY. PROBABLY BY THE DOUBLE-TIME PROTEIN (DBT). PHOSPHORYLATION COULD BE IMPLICATED IN THE STABILITY OF PER MONOMER AND IN THE FORMATION OF HETERODIMER PER-TIM (BY SIMILARITY).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.ebi.ac.uk/ebis-sib.ch/announce/> or send an email to license@ebi-sib.ch).
 CC EMBL; L06339; AAA28762.1;
 DR Flybase; Fgn0012573; Dmobjper.
 KW Biological rhythms; Repeat;
 FT NON_TER 1 1
 FT NON_TER 65 65
 FT NON_TER 65 65
 SQ SEQUENCE 65 AA; 6040 MW; 78C956E32350ED7E CRC64;

Query Match 100.0%; Score 20; DB 1; Length 65;
 Best Local Similarity 100.0%; Pred. No. 2.9e+02; Indels 0; Gaps 0;
 Matches 4; Conservative 0; Mismatches 0;

Qy 1 GSGS 4
 Db 5 GSGS 8

RESULT 7

PER_DROSA STANDARD; PRT; 66 AA.
 ID PER_DROSA
 AC Q04536;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Period circadian protein (Fragment).
 GN PER.
 OS Drosophila saltans (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Spheroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7273;
 RN [1]
 RV SEQUENCE FROM N.A.
 RP MEDLINE=93196482; PubMed=8450754;
 RA Peixoto A.A., Campos S., Costa R.H., Kyriacou C.P.;
 RT "Molecular evolution of a repetitive region within the per gene of
 RT Drosophila";
 RL Mol. Biol. Evol. 10:127-139 (1993).
 CC -|- FUNCTION: ESSENTIAL FOR BIOLOGICAL CLOCK FUNCTIONS. DETERMINES THE
 CC PERIOD LENGTH OF CIRCADIAN AND ULTRADIAN RHYTHMS; AN INCREASE IN
 CC PER DOSAGE LEADS TO SHORTENED CIRCADIAN RHYTHMS AND A DECREASE
 CC LEADS TO LENGTHENED CIRCADIAN RHYTHMS. ESSENTIAL FOR THE CIRCADIAN
 CC RHYTHMICITY OF LOCOMOTOR ACTIVITY. ECLOSION BEHAVIOR AND FOR THE
 CC RHYTHMIC COMPONENT OF THE MALE COURTSHIP SONG THAT ORIGINATES IN
 CC THE THORACIC NERVOUS SYSTEM. THE BIOLOGICAL CYCLE DEPENDS ON THE
 CC RHYTHMIC FORMATION AND NUCLEAR LOCALIZATION OF THE TIM-PER
 CC COMPLEX. LIGHT INDUCES THE DEGRADATION OF TIM, WHICH PROMOTES
 CC ELIMINATION OF PER. NUCLEAR ACTIVITY OF THE HETERODIMER
 CC COORDINATIVELY REGULATES PER AND TIM TRANSCRIPTION THROUGH A
 CC NEGATIVE FEEDBACK LOOP. BEHAVES AS A NEGATIVE ELEMENT IN CIRCADIAN
 CC TRANSCRIPTIONAL LOOP. DOES NOT APPEAR TO BIND DNA, SUGGESTING
 CC INDIRECT TRANSCRIPTIONAL INHIBITION (BY SIMILARITY).
 CC -|- SUBUNIT: FORMS HETERODIMER WITH TIMELESS (TIM); THE COMPLEX THEN
 CC TRANSLOCATES INTO THE NUCLEUS (BY SIMILARITY).
 CC -|- SUBCELLULAR LOCATION: NUCLEAR AT SPECIFIC PERIODS OF THE DAY.
 CC FIRST ACCUMULATES IN THE PERINUCLEAR REGION ABOUT ONE HOUR BEFORE
 CC TRANSLOCATION INTO THE NUCLEUS. INTERACTION WITH TIM IS REQUIRED
 CC FOR NUCLEAR LOCALIZATION (BY SIMILARITY).
 CC -|- DOMAIN: THE RUN OF GLY-THR IS IMPLICATED IN THE MAINTENANCE OF
 CC CIRCADIAN PERIOD AT DIFFERENT TEMPERATURES. DELETION OF THE REPEAT
 CC LEADS TO A SHORTENING OF THE COURTSHIP SONG CYCLE PERIOD, AND THUS
 CC COULD BE IMPORTANT FOR DETERMINING FEATURES OF SPECIES-SPECIFIC
 CC MATING BEHAVIOR (BY SIMILARITY).
 CC -|- PTM: PHOSPHORYLATED WITH A CIRCADIAN RHYTHMICITY, PROBABLY BY THE
 CC DOUBLE-TIME PROTEIN (DBT). PHOSPHORYLATION COULD BE IMPLICATED IN
 CC THE STABILITY OF PER MONOMER AND IN THE FORMATION OF HETERODIMER
 CC PER-TIM (BY SIMILARITY).

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 or send an email to license@sib-sib.ch.

EMBL; L06336; AAA28759.1; -
 DR FlyBase; Fggn0012776; Dealper.
 KW Biological rhythms; Repeat; Nuclear protein; Phosphorylation.
 FT NON_TER 1 1

FT DOMAIN 30 37 G-T REPEATS.
 FT NON_TER 66
 SQ SEQUENCE 66 AA; 5929 MW; ASCAED665177001A CRC64;

Query Match 100.0%; Score 20; DB 1; Length 66;
 Best Local Similarity 100.0%; Pred. No. 3e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GSGS 4
 Db 5 GSGS 8

RESULT 8

MCBA_ECOLI STANDARD; PRT; 69 AA.
 ID MCBA_ECOLI
 AC P05834;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Bacteriocin microcin B17 precursor (MCB17).
 GN MCBA.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RV SEQUENCE FROM N.A.
 RP MEDLINE=88217867; PubMed=3329729;
 RA Ravanino J., Herrero M., Furlong D., Moreno F., Kolter R.;
 RT "The DNA replication inhibitor microcin B17 is a
 RT forty-three-amino-acid protein containing sixty percent glycine";
 RL Proteins 1:230-238 (1986).
 RN [2]
 RV SEQUENCE FROM N.A.
 RP MEDLINE=89123111; PubMed=2844225;
 RA Genilloud O., Moreno F., Kolter R.;
 RT "DNA sequence, products, and transcriptional pattern of the genes
 RT involved in production of the DNA replication inhibitor microcin
 RT B17";
 RL J. Bacteriol. 171:1126-1135 (1989).
 RN [3]
 RV SEQUENCE OF 1-14 FROM N.A.
 RP MEDLINE=88216163; PubMed=2835580;
 RA Conell N., Han Z., Moreno F., Kolter R.;
 RT "An E. coli promoter induced by the cessation of growth";
 RL Mol. Microbiol. 1:195-201 (1987).
 RN [4]
 RV PARTIAL SEQUENCE OF 27-69.
 RP Beyer A., Stevanovic S., Freund S., Metzger J.W., Jung G.;
 RT "Isolation and structure elucidation of the 43-peptide antibiotic
 RT microcin B17";
 RL (in) Schneider C.H., Eberles A.N. (eds.);
 RL Peptides 1992, pp.117-118, ESCOM Science Publishers, Leiden (1993).
 RN [5]
 RV FUNCTION.
 RP MEDLINE=91122055; PubMed=1846808;
 RA Vizan J.L., Hernandez-Chico C., del Castillo I., Moreno F.;
 RT "The peptide antibiotic microcin B17 induces double-strand cleavage
 RT of DNA mediated by E. coli DNA gyrase";
 RL EMBO J. 10:467-476 (1991).
 RN [6]
 RV STRUCTURE BY NMR OF 1-26.
 RP MEDLINE=98213789; PubMed=9545435;
 RA Roy R.S., Kim S., Baleja J.D., Walsh C.T.;
 RT "Role of the microcin B17 propeptide in substrate recognition:
 RT solution structure and mutational analysis of MCB1-26";
 RL Chem. Biol. 5:217-228 (1998).
 CC -|- FUNCTION: THIS GLYCINE-RICH PEPTIDE ANTIBIOTIC INHIBITS DNA
 CC REPLICATION IN MANY ENTERIC BACTERIA, THAT LEADS TO INDUCTION OF
 CC THE SOS REPAIR SYSTEM, MASSIVE DNA DEGRADATION AND CELL DEATH.
 CC B17 INHIBITS TYPE II TOPOISOMERASE BY TRAPPING AN ENZYME - DNA

CC CLEAVABLE COMPLEX.
 CC -1- PTM: THE PROCESSED N-TERMINUS DOES NOT RESEMBLE A TYPICAL
 CC SECRETION SIGNAL SEQUENCE.
 CC -1- PTM: THE CYS RESIDUES AS WELL AS SOME GLY AND CYS ARE POST-
 CC TRANSLATIONALLY MODIFIED. MODIFICATIONS INCLUDE THE FORMATION OF
 CC FOUR THIAZOLE AND FOUR OXAZOLE RINGS THAT RESULT, RESPECTIVELY,
 CC FROM THE CONDENSATION OF FOUR SERINE SIDE CHAINS WITH THE CARBONYL
 CC GROUP OF THE PRECEDING AMINO ACID.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; M15469; AAA24141.1; -
 CC DR EMBL; M24253; AAA72741.1; -
 CC DR EMBL; X06417; CAA29725.1; -
 CC DR PIR; A25219; MIEC77.
 CC DR PIR; 2MLP; 22-JUL-98.
 CC KW DNA replication inhibitor; Antibiotic; Bacteriocin; Plasmid;
 CC 3D-structure.
 CC FT PROPEP 1 26
 CC FT CHAIN 27 69 BACTERIOCIN MICROCIN B17.
 CC FT DOMAIN 26 39 POLY-GLY.
 CC FT TURN 6 7
 CC FT HELIX 8 15
 CC FT TURN 16 16
 CC FT TURN 18 21
 CC FT HELIX 18 21
 CC SQ SEQUENCE 69 AA; 6013 MW; 0B1D159A832638A8 CRC64;
 CC -----
 CC Query Match 100.0%; Score 20; DB 1; Length 69;
 CC Best Local Similarity 100.0%; Pred. No. 3.1e+02;
 CC Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC -----
 CC QY 1 GSGS 4
 CC DB 64 GSGS 67
 CC -----
 CC RESULT 9
 CC MCHB ECOLI
 CC ID MCHB ECOLI STANDARD; PRT; 75 AA.
 CC AC Q9RM53;
 CC DT 16-OCT-2001 (Rel. 40, Created)
 CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Microcin H47 precursor (MccH47).
 CC GN MCHB OR C1227.
 CC OS Escherichia coli, and
 CC OS Escherichia coli O6.
 CC OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 CC OC Enterobacteriaceae; Escherichia.
 CC OX NCBI_TaxID=562, 217992;
 CC [1]
 CC RN SEQUENCE FROM N.A.
 CC RP STRAIN=H47;
 CC RC MEDLINE=99402674; PubMed=10471561;
 CC RX Rodriguez E., Gaggero C., Lavina M.,
 CC RA "The structural gene for microcin H47 encodes a peptide precursor with
 CC RT antibiotic activity."
 CC RL Antimicrob. Agents Chemother. 43:2176-2182 (1999).
 CC [2]
 CC RN SEQUENCE FROM N.A.
 CC RP STRAIN=O6:H1 / CFT073 / ATCC 700928;
 CC RC MEDLINE=22389234; PubMed=12471157;
 CC RX Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
 CC RA Rasko D., Buckles E.L., Liou S.-P., Boutin A., Hackett J., Stroud D.,
 CC RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
 CC RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
 CC RT "Extensive mosaic structure revealed by the complete genome sequence

RT of uropathogenic Escherichia coli.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
 RN [3]
 CC CHARACTERIZATION.
 CC STRAIN=H47;
 CC MEDLINE=21091907; PubMed=11181394;
 CC RX Azpiroz M.F., Rodriguez E., Lavina M.;
 CC RA "The structure, function, and origin of the microcin H47 ATP-binding
 CC cassette exporter indicate its relatedness to that of colicin V";
 CC RT Antimicrob. Agents Chemother. 45:969-972 (2001).
 CC RL Antimicrob. Agents Chemother. 45:969-972 (2001).
 CC CC -1- FUNCTION: BACTERICIDAL ANTIBIOTIC ACTIVE ON BACTERIA
 CC -1- PHYLOGENETICALLY RELATED TO THE PRODUCING STRAIN.
 CC CC -1- SUBCELLULAR LOCATION: Secreted. Probably through the mchEF abc
 CC transporter system.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; AJ009631; CAB54534.2; -
 CC DR EMBL; AE016758; AAN79685.1; ALT_INIT.
 CC KW Antibiotic; Bacteriocin; Transmembrane; Complete proteome.
 CC FT PROPEP 1 15 POTENTIAL.
 CC FT CHAIN 16 75 MICROCIN H47.
 CC FT TRANSMEM 30 50 POTENTIAL.
 CC FT SEQUENCE 75 AA; 6600 MW; 1AD5EE9F1B58123C CRC64;
 CC -----
 CC Query Match 100.0%; Score 20; DB 1; Length 75;
 CC Best Local Similarity 100.0%; Pred. No. 3.4e+02;
 CC Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC -----
 CC QY 1 GSGS 4
 CC DB 63 GSGS 66
 CC -----
 CC RESULT 10
 CC KORB STRLI
 CC ID KORB STRLI STANDARD; PRT; 80 AA.
 CC AC P22404;
 CC DT 01-AUG-1991 (Rel. 19, Created)
 CC DT 01-AUG-1991 (Rel. 19, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Korb protein.
 CC GN KORB OR COP.
 CC OS Streptomyces lividans.
 CC OG Plasmid pIJ101.
 CC OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 CC OC Streptomycineae; Streptomycetaceae; Streptomyces.
 CC OX NCBI_TaxID=1916;
 CC [1]
 CC RN SEQUENCE FROM N.A.
 CC RP STRAIN=TK24 subsp. 66;
 CC RC Radnedge L., Barallon R., Zaman S., Richards H.A., Ward J.M.;
 CC RL Submitted (JUN-1989) to the EMBL/GenBank/DDAJ databases.
 CC [2]
 CC RN SEQUENCE FROM N.A.
 CC RP MEDLINE=89008081; PubMed=3170481;
 CC RX Kendall K.J., Cohen S.N.;
 CC RA "Complete nucleotide sequence of the Streptomyces lividans plasmid
 CC RT pIJ101 and correlation of the sequence with genetic properties."
 CC RL J. Bacteriol. 170:4634-4651 (1988).
 CC CC -1- FUNCTION: REPRESSOR FOR THE TRANSCRIPTION OF CERTAIN pIJ101
 CC PROMOTERS, INCLUDING THOSE THE FROM KILA AND KILB LOC1.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X15695; CRA33730.1; -
EMBL; M21778; AA888412.1; -
PIR; S08023; S08023.
KW Transcription regulation; Plasmid; DNA-binding; Repressor.
FT DNA BIND 13 H-T-H MOTIF (BY SIMILARITY).
FT DNA BIND 56 75 H-T-H MOTIF (BY SIMILARITY).
FT CONFLICT 26 27 NV -> KL (IN REF. 2).
FT CONFLICT 37 38 DV -> EL (IN REF. 2).
SQ SEQUENCE 80 AA; 8755 MW; 9DEEB791B3C36B4D CRC64;

Query Match 100.0%; Score 20; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GSGS 4
Db 77 GSGS 80

RESULT 11
V187_BPT3
ID V187_BPT3 STANDARD; PRT; 83 AA.
AC P10302;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-MAR-1989 (Rel. 10, Last annotation update)
DE Gene 18.7 protein.
GN 18.7.
OS Bacteriophage T3.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC T7-like viruses.
OC NCBI_TaxID=10759;
OX [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=86209997; PubMed=3010556;
RA Yamada M., Fujisawa H., Kato H., Hamada K., Minagawa T.;
RT "Cloning and sequencing of the genetic right end of bacteriophage T3
DNA";
RL Virology 151:350-361(1986).

ERRATUM.
Yamada M., Fujisawa H., Kato H., Hamada K., Minagawa T.;
RA Virology 154:246-246(1986).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL; M14784; AAA92527.1; -
PIR; E23476; W8BPT3.
SQ SEQUENCE 83 AA; 9393 MW; 63128984CBAF0531 CRC64;
Query Match 100.0%; Score 20; DB 1; Length 83;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GSGS 4
Db 48 GSGS 51

RESULT 12
CRH_BACHD
ID CRH_BACHD STANDARD; PRT; 84 AA.
AC Q9K708;
Query Match 100.0%; Score 20; DB 1; Length 84;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GSGS 4
Db 48 GSGS 51

DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hpr-like protein crh (Catabolite repression Hpr).
GN CRH OR BH3566.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN SEQUENCE FROM N.A.
RP STRAIN=C-125 / JCM 9153;
RC MEDLINE=20512582; PubMed=11058132;
RX Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RA "Complete genome sequence of the alkaliphilic bacterium Bacillus
halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
CC -!- FUNCTION: Involved in carbon catabolite repression (CCR) (By
CC similarity).
CC -!- SIMILARITY: Belongs to the Hpr family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL; AP001519; BAB07285.1; -
PIR; F84095; F84095.
DR HSP; P07515; ITPF.
DR InterPro; IPR000032; Hpr_protein.
DR InterPro; IPR002114; Hpr_Serp_site.
PFam; PF00381; PTS-Hpr; 1.
PRINTS; P00107; PHOSPHOCPPHR.
PRODOM; PD002238; Hpr_protein; 1.
PROSITE; PS00589; PTS_HPR_SER; 1.
KW Phosphorylation; Complete proteome.
FT MOD_RES 46 46
FT PHOSPHORYLATION (BY HPR KINASE) (BY
FT SIMILARITY).
SQ SEQUENCE 84 AA; 9188 MW; D83BCF69C087417F CRC64;

Query Match 100.0%; Score 20; DB 1; Length 84;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GSGS 4
Db 56 GSGS 59

RESULT 13
IATP_YEAST
ID IATP_YEAST STANDARD; PRT; 85 AA.
AC P01097;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE ATPase inhibitor, mitochondrial precursor.
GN INH1 OR YDL181W OR D1305.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90202902; PubMed=2138617;
RA Ichikawa N., Yoshida Y., Hashimoto T., Ogasawara N., Yoshikawa H.,
RA Imamoto F., Tagawa K.;
RT "Activation of ATP hydrolysis by an uncoupler in mutant mitochondria
lacking an intrinsic ATPase inhibitor in yeast.";

Qy 1 GSGS 4
Db 56 GSGS 59

RL J. Biol. Chem. 265:6274-6278(1990).

RP [2] SEQUENCE FROM N.A.

RR STRAIN=8288C / FY1679;

RC MEDLINE=96021607; PubMed=8533471;

RX Verhasselt P., Voet M., Volckaert G.;

RA "New open reading frames, one of which is similar to the nifv gene of

RT Azotobacter vinelandii, found on a 12.5 kbp fragment of chromosome IV

RL of Saccharomyces cerevisiae.";

RR Yeast 11:961-966(1995).

RP [3] SEQUENCE FROM N.A.

RR Pohl T.M.;

RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.

RP [4] SEQUENCE OF 23-85.

RX MEDLINE=82075741; PubMed=6458601;

RA Matsubara H., Hase T., Hashimoto T., Tagawa K.;

RT "Amino acid sequence of an intrinsic inhibitor of mitochondrial

RL ATPase from yeast.";

RR J. Biochem. 90:1159-1165(1981).

RP [5] PARTIAL SEQUENCE OF 1-28.

RX MEDLINE=84008110; PubMed=6619114;

RA Yoshida Y., Hashimoto T., Hase T., Matsubara H., Tagawa K.;

RT "Partial amino terminal sequence of the precursor of mitochondrial

RL ATPase inhibitor protein synthesized with mRNA partially purified by

RR gel permeation chromatography.";

RP J. Biochem. 94:291-297(1983).

RP [6] MUTAGENESIS OF LYS-41.

RX MEDLINE=20169630; PubMed=10705452;

RA Ichikawa N., Fujisaka R., Kuriyama R.;

RT "Requirement for lysine-19 of the yeast mitochondrial ATPase inhibitor

RL for the stability of the inactivated inhibitor-flFo complex at higher

RR pH.";

RP Biosci. Biotechnol. Biochem. 64:89-95(2000).

CC -1- FUNCTION: FORMS A ONE-TO-ONE COMPLEX WITH ATPASE TO INHIBIT THE

CC ENZYME ACTIVITY COMPLETELY.

CC -1- SUBCELLULAR LOCATION: Mitochondrial.

CC -1- SIMILARITY: BELONGS TO THE ATPASE INHIBITOR FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; D00443; BAA00344.1; -;

DR EMBL; X83276; CAA59265.1; -;

DR EMBL; Z74229; CAA98755.1; -;

DR PIR; A35231; IWBV.

DR SGD; S0002340; INH1.

DR Pfam; PF04568; IATP; 1.

KW Mitochondrion; Transit peptide.

FT TRANSIT 1 22 MITOCHONDRION.

FT CHAIN 23 85 ATPASE INHIBITOR.

FT MUTAGEN 41 41 K-Q;E: NO LOSS OF INHIBITORY ACTIVITY.

SQ SEQUENCE 85 AA; 9870 MW; ELAF73A47F63CE58 CRC64;

Query Match 100.0%; Score 20; DB 1; Length 85;

Best Local Similarity 100.0%; Pred. No. 3.9e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSGS 4

DB 32 GSGS 35

RESULT 14

PER_DRORO

ID PER DRORO STANDARD; PRT; 86 AA.

AC 003296;

DT 01-OCT-1993 (Rel. 27, Created)

DT 01-OCT-1993 (Rel. 27, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE Period circadian protein (Fragment).

GN PER.

OS Drosophila robusta (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7257;

RP [1] SEQUENCE FROM N.A.

RX MEDLINE=93196482; PubMed=8450754;

RA Peixoto A.A., Campos S., Costa R.H., Kyriacou C.P.;

RT "Molecular evolution of a repetitive region within the per gene of

RL Drosophila.";

CC -1- FUNCTION: ESSENTIAL FOR BIOLOGICAL CLOCK FUNCTIONS. DETERMINES THE

CC PERIOD LENGTH OF CIRCADIAN AND ULTRADIAN RHYTHMS; AN INCREASE IN

CC PER DOSAGE LEADS TO SHORTENED CIRCADIAN RHYTHMS AND A DECREASE

CC LEADS TO LENGTHENED CIRCADIAN RHYTHMS. ESSENTIAL FOR THE CIRCADIAN

CC RHYTHMICITY OF LOCOMOTOR ACTIVITY, ECLOSION BEHAVIOR, AND FOR THE

CC RHYTHMIC COMPONENT OF THE MALE COURTSHIP SONG THAT ORIGINATES IN

CC THE THORACIC NERVOUS SYSTEM. THE BIOLOGICAL CYCLE DEPENDS ON THE

CC RHYTHMIC FORMATION AND NUCLEAR LOCALIZATION OF THE TIM-PER

CC COMPLEX. LIGHT INDUCES THE DEGRADATION OF TIM, WHICH PROMOTES

CC ELIMINATION OF PER. NUCLEAR ACTIVITY OF THE HETERODIMER

CC COORDINATIVELY REGULATES PER AND TIM TRANSCRIPTION THROUGH A

CC NEGATIVE FEEDBACK LOOP. BEHAVES AS A NEGATIVE ELEMENT IN CIRCADIAN

CC TRANSCRIPTIONAL LOOP. DOES NOT APPEAR TO BIND DNA, SUGGESTING

CC INDIRECT TRANSCRIPTIONAL INHIBITION (BY SIMILARITY).

CC -1- SUBUNIT: FORMS HETERODIMER WITH TIMELESS (TIM); THE COMPLEX THEN

CC TRANSLOCATES INTO THE NUCLEUS (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: NUCLEAR AT SPECIFIC PERIODS OF THE DAY.

CC FIRST ACCUMULATES IN THE PERINUCLEAR REGION ABOUT ONE HOUR BEFORE

CC TRANSLOCATION INTO THE NUCLEUS. INTERACTION WITH TIM IS REQUIRED

CC FOR NUCLEAR LOCALIZATION (BY SIMILARITY).

CC -1- DOMAIN: THE RUN OF GLY-THR IS IMPLICATED IN THE MAINTENANCE OF

CC CIRCADIAN PERIOD AT DIFFERENT TEMPERATURES. DELETION OF THE REPEAT

CC LEADS TO A SHORTENING OF THE COURTSHIP SONG CYCLE PERIOD, AND THUS

CC COULD BE IMPORTANT FOR DETERMINING FEATURES OF SPECIES-SPECIFIC

CC MATING BEHAVIOR (BY SIMILARITY).

CC -1- PTM: PHOSPHORYLATED WITH A CIRCADIAN RHYTHMICITY. PROBABLY BY THE

CC DOUBLE-TIME PROTEIN (DBT). PHOSPHORYLATION COULD BE IMPLICATED IN

CC THE STABILITY OF PER MONOMER AND IN THE FORMATION OF HETERODIMER

CC PER-TIM (BY SIMILARITY).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; L06340; AAA28763.1; -;

DR Flybase; FBgn0012755; Drob\per.

KW Biological rhythms; Repeat; Nuclear protein; Phosphorylation.

FT NON TER 1 1

FT DOMAIN 30 53 G-T REPEATS.

FT NON TER 86 86

SQ SEQUENCE 86 AA; 7746 MW; C374AD7D7A06FOCE CRC64;

Query Match 100.0%; Score 20; DB 1; Length 86;

Best Local Similarity 100.0%; Pred. No. 3.9e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSGS 4

DB 5 GSGS 8

```
RESULT 15
RK27_GUITH
ID RK27_GUITH STANDARD; PRT; 86 AA.
AC O78430;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chloroplast 50S ribosomal protein L27.
GN RP127
OS Guillardia theta (Cryptomonas phi).
OG Chloroplast.
OC Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.
OX NCBI_TaxID=55529;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93128221; PubMed=9929392;
RA Douglas S.E., Penny S.L.;
RT "The plastid genome of the cryptophyte alga, Guillardia theta:
RT complete sequence and conserved syntenic groups confirm its common
RT ancestry with red algae."
RL J. Mol. Evol. 48:236-244(1999).
CC -!- SIMILARITY: BELONGS TO THE L27P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch)
CC -----
CC EMBL; AF041468; AAC35615.1; -.
DR HAMAP; MF_00539; -.
DR InterPro; IPR001684; Ribosomal_L27.
DR Pfam; PF01016; Ribosomal_L27; 1.
DR PRINTS; PR00063; RIBOSOMALL27.
DR ProDom; PD003114; Ribosomal_L27; 1.
DR TIGRFAMs; TIGR00062; L27; 1_L27; 1.
DR PROSITE; PS00831; RIBOSOMAL_L27; 1.
KW Ribosomal protein; Chloroplast.
SQ SEQUENCE 86 AA; 9534 MW; 08BEC51FC44389A8 CRC64;

Query Match 100.0%; Score 20; DB 1; Length 86;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSGS 4
   |||
Db 6 GSGS 9
```

Search completed: February 11, 2004, 22:18:36

Job time : 34 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 21:41:00 ; Search time 87 Seconds

(without alignments)
11.964 Million cell updates/sec

Title: US-09-813-341-10

Perfect score: 20

Sequence: 1 GSGS 4

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	10	11 Q8BHN2	Q8BHN2 mus musculus
2	20	100.0	14	11 Q8CJA8	Q8CJA8 mus musculus
3	20	100.0	15	11 Q8CJA9	Q8CJA9 mus musculus
4	20	100.0	29	4 Q9BSQ3	Q9BSQ3 homo sapien
5	20	100.0	29	11 Q98YV5	Q98YV5 mus musculus
6	20	100.0	30	11 Q923S3	Q923S3 mus musculus
7	20	100.0	31	5 Q23476	Q23476 caenorhabdi
8	20	100.0	32	16 Q9A506	Q9A506 caulobacter
9	20	100.0	35	16 Q8F0E7	Q8F0E7 leptospira
10	20	100.0	38	11 Q99N61	Q99N61 rattus norv
11	20	100.0	40	6 Q9BDE8	Q9BDE8 sus scrofa
12	20	100.0	40	6 Q18922	Q18922 sus scrofa
13	20	100.0	41	10 Q9S9F1	Q9S9F1 brassica na
14	20	100.0	41	11 Q8BQ53	Q8BQ53 mus musculus
15	20	100.0	42	12 Q90060	Q90060 zucchini ye
16	20	100.0	42	16 Q98Q40	Q98Q40 mycoplasma

17	20	100.0	43	10 Q80797	Q80797 arabidopsis
18	20	100.0	43	13 Q9PSF3	Q9PSF3 gallus gall
19	20	100.0	46	2 Q93GL0	Q93GL0 mycobacteri
20	20	100.0	46	2 Q9APK7	Q9APK7 streptococc
21	20	100.0	46	16 Q99QI5	Q99QI5 streptococc
22	20	100.0	49	16 Q8FZB4	Q8FZB4 bruceella su
23	20	100.0	50	12 Q8QUM7	Q8QUM7 infectious
24	20	100.0	50	16 Q8VJX1	Q8VJX1 mycobacteri
25	20	100.0	51	2 Q9F0Y4	Q9F0Y4 streptomyce
26	20	100.0	51	2 Q9ZHL5	Q9ZHL5 streptococc
27	20	100.0	51	2 Q9XB45	Q9XB45 escherichia
28	20	100.0	51	11 Q8CIW7	Q8CIW7 mus musculu
29	20	100.0	51	12 Q8JTI1	Q8JTI1 little cher
30	20	100.0	51	16 Q8PAS0	Q8PAS0 xanthomonas
31	20	100.0	51	16 Q8NLJ2	Q8NLJ2 corynebacte
32	20	100.0	51	16 Q8PE64	Q8PE64 escherichia
33	20	100.0	52	6 Q77483	Q77483 canis fami
34	20	100.0	52	10 Q8S2Q9	Q8S2Q9 oryza sativ
35	20	100.0	53	2 Q93SL5	Q93SL5 streptococc
36	20	100.0	53	6 Q28023	Q28023 bos taurus
37	20	100.0	53	6 Q19022	Q19022 mirounga le
38	20	100.0	53	6 Q28551	Q28551 ovis aries
39	20	100.0	53	6 Q28981	Q28981 sus scrofa
40	20	100.0	53	10 Q8W2Y8	Q8W2Y8 oryza sativ
41	20	100.0	53	16 Q69194	Q69194 streptococc
42	20	100.0	54	2 Q8GP18	Q8GP18 streptococc
43	20	100.0	54	5 Q9U7L2	Q9U7L2 plasmodium
44	20	100.0	54	16 Q8YQF2	Q8YQF2 anabaena sp
45	20	100.0	55	10 Q8LRE2	Q8LRE2 oryza sativ

ALIGNMENTS

RESULT 1

Q8BHN2 PRELIMINARY; PRT; 10 AA.

AC Q8BHN2; DT 01-MAR-2003 (TRENBLrel. 23, Created)

DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)

DE Methylenetetrahydrofolate reductase short isoform (Fragment).

GN MTHFR.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=129/Sv. and BALB/c;

RX MEDLINE=2257759; PubMed=12370778;

RA Tran P., Leclerc D., Chan M., Pal A., Hlou-Tim F., Wu Q., Goyette P.,

RA Artigas C., Milos R., Rozen R.;

RT "Multiple transcription start sites and alternative splicing in the

RT methylenetetrahydrofolate reductase gene result in two enzyme

RT isoforms.";

RL Mamm. Genome 13:483-492(2002).

DR EMBL; AY046557; AAL17641.1; -.

DR EMBL; AF398931; AAN40867.1; -.

FT NON_TER 10

SQ SEQUENCE 10 AA; 1007 MW; 01695CB8640DD814 CRC64;

Query Match 100.0%; Score 20; DB 11; Length 10;

Best Local Similarity 100.0%; Pred. No. 3.6e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSGS 4

Db 7 GSGS 10

RESULT 2

Q8CJA8

```

ID Q8CJA8 PRELIMINARY; PRT; 14 AA.
AC Q8CJA8;
DT 01-WAR-2003 (TrEMBLrel. 23, Created)
DT 01-WAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-WAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Methylene tetrahydrofolate reductase short isoform (Fragment).
GN MTHFR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV;
RX MEDLINE=22257759; PubMed=12370778;
RA Tran P., Leclerc D., Chan M., Pai A., Hiou-Tim F., Wu Q., Goyette P.,
RA Artigas C., Milos R., Rozen R.;
RT "Multiple transcription start sites and alternative splicing in the
RT methylenetetrahydrofolate reductase gene result in two enzyme
RT isoforms.";
RL Mamm. Genome 13:483-492(2002).
RL EMBL; AF404271; AAN40873.1; -.
FT NON TER 14 14
SQ SEQUENCE 14 AA; 1472 MW; 28DD341AC1695CB8 CRC64;

Query Match 100.0%; Score 20; DB 11; Length 14;
Best Local Similarity 100.0%; Pred. No. 5e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 0;

Qy 1 GSGS 4
Db 7 GSGS 10

RESULT 3
Q8CJA9 PRELIMINARY; PRT; 15 AA.
AC Q8CJA9;
DT 01-WAR-2003 (TrEMBLrel. 23, Created)
DT 01-WAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-WAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Methylene tetrahydrofolate reductase short isoform (Fragment).
GN MTHFR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV;
RX MEDLINE=22257759; PubMed=12370778;
RA Tran P., Leclerc D., Chan M., Pai A., Hiou-Tim F., Wu Q., Goyette P.,
RA Artigas C., Milos R., Rozen R.;
RT "Multiple transcription start sites and alternative splicing in the
RT methylenetetrahydrofolate reductase gene result in two enzyme
RT isoforms.";
RL Mamm. Genome 13:483-492(2002).
RL EMBL; AF404270; AAN40872.1; -.
FT NON TER 15 15
SQ SEQUENCE 15 AA; 1559 MW; 8D78DD341AC1695C CRC64;

Query Match 100.0%; Score 20; DB 11; Length 15;
Best Local Similarity 100.0%; Pred. No. 5.4e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 0;

Qy 1 GSGS 4
Db 7 GSGS 10

RESULT 4
Q8BSQ3 PRELIMINARY; PRT; 29 AA.
ID Q8BSQ3

```

```

AC Q9BSQ3;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC004906; AA04906.1; -.
SQ SEQUENCE 29 AA; 2926 MW; 8C871B5287F2D6AA CRC64;

Query Match 100.0%; Score 20; DB 4; Length 29;
Best Local Similarity 100.0%; Pred. No. 1.1e+03; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 0;

Qy 1 GSGS 4
Db 19 GSGS 22

RESULT 5
Q99JY5 PRELIMINARY; PRT; 29 AA.
ID Q99JY5
AC Q99JY5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to protein kinase, DNA activated, catalytic polypeptide
DE interacting protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC005570; AA05570.1; -.
DR HSSP; Q99828; IDGV.
KW Kinase.
SQ SEQUENCE 29 AA; 3242 MW; E19D6177AF420612 CRC64;

Query Match 100.0%; Score 20; DB 11; Length 29;
Best Local Similarity 100.0%; Pred. No. 1.1e+03; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 0;

Qy 1 GSGS 4
Db 3 GSGS 6

RESULT 6
Q923S3 PRELIMINARY; PRT; 30 AA.
ID Q923S3
AC Q923S3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Nonclathrin coat protein gamma2-COP (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129;

Query Match 100.0%; Score 20; DB 11; Length 29;
Best Local Similarity 100.0%; Pred. No. 1.1e+03; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 0;

Qy 1 GSGS 4
Db 3 GSGS 6

RESULT 6
Q923S3 PRELIMINARY; PRT; 30 AA.
ID Q923S3
AC Q923S3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Nonclathrin coat protein gamma2-COP (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129;

```

```

RA Yun J., Lee Y.J., Park C.W., Seong J.-K., Kim H., Chung J.H.;
RT "Methylation status at the promoter-associate CpG island of Copg2 and
RT existence of a novel antisense.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF402597; AAK94484.1; -.
KW Coat protein. 30
FT NON_TER 30
SQ SEQUENCE 30 AA; 3408 MW; DE99F1CADE0475FE CRC64;

Query Match 100.0%; Score 20; DB 11; Length 30;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSGS 4
DB 13 GSGS 16

RESULT 7
Q23476 PRELIMINARY; PRT; 31 AA.
AC Q23476;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical 3.3 kDa protein.
GN ZK402.4
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Favello A.;
RT "The sequence of C. elegans cosmid ZK402.";
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U40956; AAA81755.1; -.
DR WormPep; ZK402.4; CE05076.
KW Hypothetical protein.
SQ SEQUENCE 31 AA; 3271 MW; 5944F94215F25689 CRC64;

Query Match 100.0%; Score 20; DB 5; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSGS 4
DB 24 GSGS 27

RESULT 8
Q9A506 PRELIMINARY; PRT; 32 AA.
AC Q9A506;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein CC2665.

```

```

GN CC2665.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CBI5;
RX MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Sait J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Utterback T., Iran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL; AE005934; AAK24632.1; -.
DR TIGR; CC2665; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 32 AA; 3533 MW; 36D5E7CF17F98513 CRC64;

Query Match 100.0%; Score 20; DB 16; Length 32;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSGS 4
DB 25 GSGS 28

RESULT 9
Q8F0E7 PRELIMINARY; PRT; 35 AA.
AC Q8F0E7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN LA3548.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA Ren S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE011511; AAN50746.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 35 AA; 4241 MW; EF45ED8DF96940C8 CRC64;

Query Match 100.0%; Score 20; DB 16; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSGS 4
DB 13 GSGS 16

RESULT 10
Q99N61 PRELIMINARY; PRT; 38 AA.
AC Q99N61;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Neuroglobin-2.
GN NGB-2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Wang H., Gao X., Wang B., Huang Y., Han J.;
RT "Rat neuroglobin-2 (NGB-2) cDNA, partial sequence.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB056656; BAB39150.1;
SQ SEQUENCE 38 AA; 3868 MW; CD204E7C116E5E04 CRC64;

Query Match 100.0%; Score 20; DB 11; Length 38;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSGS 4
DB 26 GSGS 29

RESULT 11
Q9BDE8 PRELIMINARY; PRT; 40 AA.
AC Q9BDE8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Heat shock protein 70.2 (Fragment).
GN HSP70.2.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=92175874; PubMed=1339404;
RA Peelman L.J., Van de Weghe A.R., Coppieiers W.R., Van Zevenen A.J.,
RT Bouquet Y.H.;
RT "Complete nucleotide sequence of a porcine HSP70 gene.";
RL Immunogenetics 35:286-289(1992).
RN [2]
RP SEQUENCE FROM N.A.
RA Scherwin M., Maak S., Hagendorf A., von Lengerken G., Seyfert H.M.;
RT "A 3'-UTR variant of the inducible porcine hsp70.2 gene affects mRNA
RT stability.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ310378; CAC36994.1;
FT NON_TER 1
FT CONFLICT 18 18 A -> P (IN REF. 1).
FT CONFLICT 24 24 Q -> P (IN REF. 1).
FT CONFLICT 25 25 A -> D (IN REF. 1).
FT CONFLICT 26 26 P -> L (IN REF. 1).
SQ SEQUENCE 40 AA; 3674 MW; BOCF5E206E3AEBF8 CRC64;

Query Match 100.0%; Score 20; DB 6; Length 40;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSGS 4
DB 29 GSGS 32

RESULT 12
O18922 PRELIMINARY; PRT; 40 AA.
AC O18922;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Epithelial-cadherin (E-cadherin) (Uvomorulin) (CaM 120/80) (Fragment).
GN CAD1.
OS Sus scrofa (Pig).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=RETINA;
RA Lutz D.A., Zheng J.J.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.
CC -!- FUNCTION: CADHERINS INTERACT WITH THEMSELVES IN A HOMOPHILIC
CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE
CC SORTING OF HETEROGENEOUS CELL TYPES. E-CADHERIN HAS A POTENT
CC INVASIVE SUPPRESSOR ROLE.
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE CADHERIN FAMILY.
DR EMBL: AF033019; AAB87474.1;
DR InterPro: IPR002126; Cadherin.
DR InterPro: IPR000233; Cadherin_C_term.
DR Pfam: PF01649; Cadherin_C_term_1.
DR PROSITE: PS00232; CADHERIN_1; PARTIAL.
KW Cell adhesion; Glycoprotein; Phosphorylation; Transmembrane;
KW Calcium-binding; Repeat.
FT NON_TER 1
FT CHAIN <1 >40 EPITHELIAL-CADHERIN.
FT DOMAIN <1 >40 CYTOPLASMIC (POTENTIAL).
FT NON_TER 40
SQ SEQUENCE 40 AA; 4449 MW; CB928720A51B8372 CRC64;

Query Match 100.0%; Score 20; DB 6; Length 40;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSGS 4
DB 9 GSGS 12

RESULT 13
Q9S9F1 PRELIMINARY; PRT; 41 AA.
AC Q9S9F1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE Napin short chain S4=CALMODULIN antagonist/calcium-dependent protein
DE kinase substrate.
OS Brassica napus (Rape).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3708;
RN [1]
RP SEQUENCE.
RX MEDLINE=96283790; PubMed=8679670;
RA Neumann G.M., Condron R., Thomas I., Polya G.M.;
RT "Purification and sequencing of multiple forms of Brassica napus seed
RT napin small chains that are calmodulin antagonists and substrates for
RT plant calcium-dependent protein kinase.";
RL Biochim. Biophys. Acta 1295:23-33(1996).
SQ SEQUENCE 41 AA; 4652 MW; 8BC6738503380553 CRC64;

Query Match 100.0%; Score 20; DB 10; Length 41;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSGS 4
DB 35 GSGS 38

RESULT 14
Q9BQ53

```



```
ID QBQ53 PRELIMINARY; PRT; 41 AA.
AC QBQ53;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE SRY-box containing gene 4 (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Dorsal root ganglion;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK051540; BAC34667.1; -.
FT NON TER 1
SQ SEQUENCE 41 AA; 4667 MW; C7D0B495BDBDFE65 CRC64;

Query Match 100.0%; Score 20; DB 11; Length 41;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSGS 4
Db |||||
6 GSGS 9

RESULT 15
Q90060
ID Q90060 PRELIMINARY; PRT; 42 AA.
AC Q90060;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Coat protein (Fragment).
GN COAT PROTEIN, CP.
OS Zucchini yellow mosaic virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Potyvirus.
OX NCBI_TaxID=12232;
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=93019038; PubMed=1402810;
RX Gal-On A., Antignus Y., Rosner A., Raccach B.;
RA "A zucchini yellow mosaic virus coat protein gene mutation restores
RT aphid transmissibility but has no effect on multiplication."
RL J. Gen. Virol. 73:2183-2187 (1992).
DR EMBL; S46009; AAB2350.1; -.
FT NON TER 1
SQ SEQUENCE 42 AA; 4209 MW; 587AD79DE4A1711B CRC64;

Query Match 100.0%; Score 20; DB 12; Length 42;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSGS 4
Db |||||
29 GSGS 32

Search completed: February 11, 2004, 22:20:14
Job time : 93 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 22:16:34 ; Search time 30 Seconds
(without alignments)
5.641 Million cell updates/sec

Title: US-09-813-341-10
Perfect score: 20
Sequence: 1 GSGS 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/prodata/2/iaa/5A COMB.pep.*
2: /cgn2_6/prodata/2/iaa/5B COMB.pep.*
3: /cgn2_6/prodata/2/iaa/6A COMB.pep.*
4: /cgn2_6/prodata/2/iaa/6B COMB.pep.*
5: /cgn2_6/prodata/2/iaa/PCTUS COMB.pep.*
6: /cgn2_6/prodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	4	2	US-08-482-651-5
2	20	100.0	4	3	US-08-602-999A-104
3	20	100.0	4	3	US-09-113-921-90
4	20	100.0	4	3	US-09-113-921-91
5	20	100.0	4	3	US-08-660-092-8
6	20	100.0	4	3	US-09-198-723A-21
7	20	100.0	4	3	US-08-918-288-80
8	20	100.0	4	3	US-09-282-357-80
9	20	100.0	4	4	US-08-278-865-104
10	20	100.0	4	4	US-09-344-456-4
11	20	100.0	4	4	US-09-561-366B-39
12	20	100.0	4	4	US-08-160-513-8
13	20	100.0	4	4	US-09-500-124-104
14	20	100.0	4	4	US-10-114-176-39
15	20	100.0	4	4	US-09-451-067-90
16	20	100.0	4	4	US-09-451-067-91
17	20	100.0	4	4	US-08-347-335A-8
18	20	100.0	5	2	US-09-194-613-11
19	20	100.0	6	1	US-08-264-002-15
20	20	100.0	6	2	US-08-803-899-21
21	20	100.0	6	3	US-08-918-288-81
22	20	100.0	6	3	US-09-282-357-81
23	20	100.0	6	4	US-09-235-230-31
24	20	100.0	8	2	US-08-529-190B-18
25	20	100.0	8	2	US-08-803-899-18
26	20	100.0	8	2	US-08-830-853-25
27	20	100.0	8	3	US-09-029-424-19

28	20	100.0	8	3	US-08-890-615-9	Sequence 9, Appli
29	20	100.0	8	3	US-08-918-288-52	Sequence 52, Appl
30	20	100.0	8	3	US-09-282-357-52	Sequence 52, Appl
31	20	100.0	8	4	US-09-246-290A-9	Sequence 9, Appli
32	20	100.0	9	3	US-08-808-599A-26	Sequence 26, Appl
33	20	100.0	9	4	US-09-593-870A-39	Sequence 39, Appl
34	20	100.0	10	2	US-08-701-124-72	Sequence 72, Appl
35	20	100.0	10	2	US-08-803-899-26	Sequence 26, Appl
36	20	100.0	10	2	US-08-350-260A-437	Sequence 437, App
37	20	100.0	10	2	US-08-253-678A-19	Sequence 19, Appl
38	20	100.0	10	3	US-09-130-225-72	Sequence 72, Appl
39	20	100.0	10	3	US-08-582-134B-19	Sequence 19, Appl
40	20	100.0	10	3	US-08-918-288-82	Sequence 82, Appl
41	20	100.0	10	3	US-09-282-357-82	Sequence 82, Appl
42	20	100.0	10	4	US-09-230-548-3	Sequence 3, Appli
43	20	100.0	10	4	US-09-455-061-72	Sequence 72, Appl
44	20	100.0	10	4	US-09-104-337A-437	Sequence 437, App
45	20	100.0	11	1	US-08-200-900A-37	Sequence 37, Appl

ALIGNMENTS

RESULT 1
US-08-482-651-5
; Sequence 5, Application US/08482651
; Patent No. 5874409
; GENERAL INFORMATION:
; APPLICANT: Victoria, Edward J.
; TITLE OF INVENTION: aPL IMMUNOREACTIVE PEPTIDES, CONJUGATES
; TITLE OF INVENTION: THEROF AND METHODS OF TREATMENT FOR aPL ANTIBODY-MEDIATED
; TITLE OF INVENTION: PATHOLOGIES
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,651
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Park, Freddie K.
; REGISTRATION NUMBER: 35,636
; REFERENCE/DOCKET NUMBER: 25231-20061.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-482-651-5

Query Match 100.0%; Score 20; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSGS 4
DB 1 GSGS 4

RESULT 2
US-08-602-999A-104
; Sequence 104, Application US/08602999A
; Patent No. 6184205
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: KAY, Brian K.
; APPLICANT: THORN, Judith M.
; APPLICANT: QUILLIAM, Lawrence A.
; APPLICANT: DER, Channing J.
; APPLICANT: FOWLKES, Dana M.
; APPLICANT: RIDER, James E.
; TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
; TITLE OF INVENTION: ISOLATING AND USING SAME
; NUMBER OF SEQUENCES: 467
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/602,999A
; FILING DATE: 16-FEB-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-202
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 104:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-602-999A-104

Query Match 100.0%; Score 20; DB 3; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GSGS 4
Db 1 GSGS 4

RESULT 3
US-09-113-921-90
; Sequence 90, Application US/09113921
; Patent No. 6193981
; GENERAL INFORMATION:
; APPLICANT: Goldstein, Gideon
; TITLE OF INVENTION: Methods and Compositions for Impairing
; TITLE OF INVENTION: Multiplication of HIV-1
; NUMBER OF SEQUENCES: 124
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr., P.O. Box 457
; CITY: Spring House
; STATE: PA
; COUNTRY: USA

ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/113,921
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/893,853
; FILING DATE: 11-JUL-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: GGP2AUSA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 90:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-113-921-90
Query Match 100.0%; Score 20; DB 3; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GSGS 4
Db 1 GSGS 4
RESULT 4
US-09-113-921-91
; Sequence 91, Application US/09113921
; Patent No. 6193981
; GENERAL INFORMATION:
; APPLICANT: Goldstein, Gideon
; TITLE OF INVENTION: Methods and Compositions for Impairing
; TITLE OF INVENTION: Multiplication of HIV-1
; NUMBER OF SEQUENCES: 124
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr., P.O. Box 457
; CITY: Spring House
; STATE: PA
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/113,921
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/893,853
; FILING DATE: 11-JUL-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: GGP2AUSA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818

INFORMATION FOR SEQ ID NO: 91:

SEQUENCE CHARACTERISTICS:

LENGTH: 4 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: peptide

FEATURE:

NAME/KEY: Modified-site

LOCATION: 4

OTHER INFORMATION: /note= "Biccytin-amide is attached

OTHER INFORMATION: to Ser in position 4."

US-09-113-921-91

Query Match 100.0%; Score 20; DB 3; Length 4;

Best Local Similarity 100.0%; Pred. No. 2.5e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GSGS 4

Db 1 GSGS 4

RESULT 5

US-08-660-092-8

Sequence 8, Application US/08660092

Patent No. 6207160

GENERAL INFORMATION:

APPLICANT: Victoria, Edward J.

APPLICANT: Marquis, David M.

APPLICANT: Jones, David S.

APPLICANT: Yu, Lin

TITLE OF INVENTION: aPL IMMUNOREACTIVE PEPTIDES, CONJUGATES

TITLE OF INVENTION: THEREOF AND METHODS OF TREATMENT FOR aPL ANTIBODY-MEDIATED

TITLE OF INVENTION: PATHOLOGIES

NUMBER OF SEQUENCES: 216

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORRISON & FOERSTER

STREET: 755 PAGE MILL ROAD

CITY: PALO ALTO

STATE: CA

COUNTRY: USA

ZIP: 94304-1018

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/660,092

FILING DATE: 06-JUN-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Park, Freddie K.

REGISTRATION NUMBER: 35,636

REFERENCE/DOCKET NUMBER: 25231-20061.20

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 813-5600

TELEFAX: (415) 494-0792

TELEX: 706141

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 4 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-660-092-8

Query Match 100.0%; Score 20; DB 3; Length 4;

Best Local Similarity 100.0%; Pred. No. 2.5e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GSGS 4

Db 1 GSGS 4

RESULT 6

US-09-198-723A-21

Sequence 21, Application US/09198723A

Patent No. 6211338

GENERAL INFORMATION:

APPLICANT: Malcolm, Bruce

APPLICANT: Taremi, Shahril S.

APPLICANT: Weber, Patricia

APPLICANT: Yao, Nanhua

TITLE OF INVENTION: Covalent Complexes of Hepatitis C Virus

TITLE OF INVENTION: NS3 protease and NS4A Cofactor Peptide

NUMBER OF SEQUENCES: 123

CORRESPONDENCE ADDRESS:

ADDRESSEE: Schering-Plough Corp.

STREET: 2000 Galloping Hill Road

CITY: Kenilworth

STATE: New Jersey

COUNTRY: USA

ZIP: 07030

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: Power Macintosh

OPERATING SYSTEM: 8.0.1

SOFTWARE: Microsoft Word 6.0.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/198,723A

FILING DATE: 24 NOV 1998

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: McLaughlin, Jaye P.

REGISTRATION NUMBER: 41,211

REFERENCE/DOCKET NUMBER: JB0800

TELECOMMUNICATION INFORMATION:

TELEPHONE: (908)298-5056

TELEFAX: (908)298-5388

INFORMATION FOR SEQ ID NO: 21:

SEQUENCE CHARACTERISTICS:

LENGTH: 4 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-09-198-723A-21

Query Match 100.0%; Score 20; DB 3; Length 4;

Best Local Similarity 100.0%; Pred. No. 2.5e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GSGS 4

Db 1 GSGS 4

RESULT 7

US-08-918-288-80

Sequence 80, Application US/08918288

Patent No. 6238890

GENERAL INFORMATION:

APPLICANT: BOIME, Irving

APPLICANT: MOYLE, William R.

TITLE OF INVENTION: SINGLE-CHAIN FORMS OF THE

TITLE OF INVENTION: GLYCOPROTEIN HORMONE QUARTET

NUMBER OF SEQUENCES: 83

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORRISON & FOERSTER

STREET: 2000 Pennsylvania Avenue, NW, suite 5500

CITY: Washington

STATE: DC

COUNTRY: USA


```

Query Match      100.0%; Score 20; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSGS 4
Db 1 GSGS 4

RESULT 10
US-09-344-456-4
; Sequence 4, Application US/09344456A
; Patent No. 6326137
; GENERAL INFORMATION:
; APPLICANT: Hong, Zhi
; APPLICANT: Lai, Vicky C.H.
; APPLICANT: Lau, Johnson Y.N.
; TITLE OF INVENTION: HEPATITIS C VIRUS PROTEASE-DEPENDENT CHIMERIC
; FILE REFERENCE: IN01038
; CURRENT APPLICATION NUMBER: US/09/344,456A
; CURRENT FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 4
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Linker
US-09-344-456-4

Query Match      100.0%; Score 20; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSGS 4
Db 1 GSGS 4

RESULT 11
US-09-561-366B-39
; Sequence 39, Application US/09561366B
; Patent No. 6399067
; GENERAL INFORMATION:
; APPLICANT: Goldstein, Gideon
; TITLE OF INVENTION: Methods and Compositions for Impairing Multiplication of HIV-1
; FILE REFERENCE: GGP3USA
; CURRENT APPLICATION NUMBER: US/09/561,366B
; CURRENT FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 39
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-09-561-366B-39

Query Match      100.0%; Score 20; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSGS 4
Db 1 GSGS 4

RESULT 12
US-09-160-513-8
; Sequence 8, Application US/09160513
; Patent No. 6410775
; GENERAL INFORMATION:

```

```

; APPLICANT: Victoria, Edward J.
; APPLICANT: Marquis, David M.
; APPLICANT: Jones, David S.
; APPLICANT: Yu, Lin
; TITLE OF INVENTION: APL IMMUNOREACTIVE PEPTIDES, CONJUGATES THEREOF AND METHODS O
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/09/160,513
; FILING DATE: 1998-DEC-24
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: CATHERINE M. POLIZZI
; REGISTRATION NUMBER: 40,130
; REFERENCE/DOCKET NUMBER: 25231-20061.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 813-5600
; TELEFAX: (650) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-160-513-8

Query Match      100.0%; Score 20; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSGS 4
Db 1 GSGS 4

RESULT 13
US-09-500-124-104
; Sequence 104, Application US/09500124
; Patent No. 6432920
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: KAY, Brian K.
; APPLICANT: THORN, Judith M.
; APPLICANT: QUILLIAM, Lawrence A.
; APPLICANT: DER, Channing J.
; APPLICANT: FOWLKES, Dana M.
; APPLICANT: RIDER, James E.
; TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
; TITLE OF INVENTION: ISOLATING AND USING SAME
; NUMBER OF SEQUENCES: 467
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/500,124
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER: 08/602,999
; APPLICATION NUMBER: 08/602,999
; FILING DATE: 16-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistock, S Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-202
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 104:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-09-500-124-104

Query Match      100.0%; Score 20; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.5e+05; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GSGS 4
Db      1 GSGS 4

RESULT 14
US-10-114-176-39
; Sequence 39, Application US/10114176
; Patent No. 6524582
; GENERAL INFORMATION:
; APPLICANT: Goldstein, Gideon
; TITLE OF INVENTION: Methods and Compositions for Impairing Multiplication of HIV-1
; FILE REFERENCE: GGP2USA
; CURRENT APPLICATION NUMBER: US/10/114,176
; CURRENT FILING DATE: 2002-04-02
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 39
; LENGTH: 4
; TYPE: PPT
; ORGANISM: Human immunodeficiency virus type 1
US-10-114-176-39

Query Match      100.0%; Score 20; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.5e+05; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GSGS 4
Db      1 GSGS 4

RESULT 15
US-09-451-067-90
; Sequence 90, Application US/09451067
; Patent No. 6525179
; GENERAL INFORMATION:
; APPLICANT: Goldstein, Gideon
; TITLE OF INVENTION: Methods and Compositions for Impairing
; TITLE OF INVENTION: Multiplication of HIV-1
; NUMBER OF SEQUENCES: 124
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr., P.O. Box 457
; CITY: Spring House
```

```
; STATE: PA
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/451,067
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/113,921
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: GGP2AUSA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 90:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-451-067-90

Query Match      100.0%; Score 20; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.5e+05; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GSGS 4
Db      1 GSGS 4

Search completed: February 11, 2004, 22:23:20
Job time : 32 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 22:18:03 ; Search time 76 Seconds
(without alignments)
11.020 Million cell updates/sec

Title: US-09-813-341-10
Perfect score: 20
Sequence: 1 GSGS 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	4	9	US-09-938-315-104
2	20	100.0	4	12	US-09-949-039-54
3	20	100.0	4	12	US-10-262-435-90
4	20	100.0	4	12	US-10-262-435-91
5	20	100.0	4	12	US-10-323-013-39
6	20	100.0	4	12	US-10-161-791-104
7	20	100.0	4	12	US-10-086-208-90
8	20	100.0	4	12	US-10-086-208-91
9	20	100.0	4	12	US-10-044-844-8
10	20	100.0	4	14	US-10-114-176-39
11	20	100.0	4	15	US-10-133-133A-7
12	20	100.0	5	15	US-10-099-424B-8
13	20	100.0	5	15	US-10-274-638-6
14	20	100.0	7	12	US-10-412-382-17
15	20	100.0	8	15	US-10-104-939-48

16	20	100.0	9	10	US-09-780-668A-24	Sequence 24, Appl
17	20	100.0	9	12	US-10-307-389-47	Sequence 47, Appl
18	20	100.0	10	10	US-09-969-192-72	Sequence 72, Appl
19	20	100.0	10	11	US-09-572-404B-152	Sequence 152, App
20	20	100.0	10	11	US-09-572-404B-304	Sequence 204, App
21	20	100.0	10	11	US-09-572-404B-206	Sequence 206, App
22	20	100.0	10	11	US-09-572-404B-356	Sequence 256, App
23	20	100.0	10	11	US-09-572-404B-316	Sequence 316, App
24	20	100.0	10	11	US-09-572-404B-318	Sequence 318, App
25	20	100.0	10	11	US-09-572-404B-322	Sequence 322, App
26	20	100.0	10	11	US-09-572-404B-453	Sequence 453, App
27	20	100.0	10	11	US-09-572-404B-455	Sequence 455, App
28	20	100.0	10	11	US-09-572-404B-457	Sequence 457, App
29	20	100.0	10	11	US-09-572-404B-459	Sequence 459, App
30	20	100.0	10	11	US-09-572-404B-471	Sequence 471, App
31	20	100.0	10	11	US-09-572-404B-473	Sequence 473, App
32	20	100.0	10	11	US-09-572-404B-475	Sequence 475, App
33	20	100.0	10	11	US-09-572-404B-493	Sequence 493, App
34	20	100.0	10	11	US-09-572-404B-495	Sequence 495, App
35	20	100.0	10	11	US-09-572-404B-497	Sequence 497, App
36	20	100.0	10	11	US-09-572-404B-499	Sequence 499, App
37	20	100.0	10	11	US-09-572-404B-501	Sequence 501, App
38	20	100.0	10	11	US-09-572-404B-503	Sequence 503, App
39	20	100.0	10	11	US-09-572-404B-505	Sequence 505, App
40	20	100.0	10	11	US-09-572-404B-507	Sequence 507, App
41	20	100.0	10	11	US-09-572-404B-509	Sequence 509, App
42	20	100.0	10	11	US-09-572-404B-511	Sequence 511, App
43	20	100.0	10	11	US-09-572-404B-517	Sequence 517, App
44	20	100.0	10	11	US-09-572-404B-518	Sequence 518, App
45	20	100.0	10	11	US-09-572-404B-519	Sequence 519, App

ALIGNMENTS

RESULT 1

US-09-938-315-104

; Sequence 104, Application US/09938315

; Patent No. US20020091085A1

; GENERAL INFORMATION:

; APPLICANT: KAY, BRIAN K.

; SPARKS, ANDREW B.

; THORN, JUDITH M.

; QUILLIAM, LAWRENCE A.

; DER, CHANNING J.

; TITLE OF INVENTION: SEC SH3 BINDING PEPTIDES AND METHODS OF

; ISOLATING AND USING SAME

; NUMBER OF SEQUENCES: 106

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,

; P.C.

; STREET: 1755 S. Jefferson Davis Highway, Suite 400

; CITY: Arlington

; STATE: Virginia

; COUNTRY: U.S.A.

; ZIP: 22202

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/938,315

; FILING DATE: 23-Aug-2001

; CLASSIFICATION: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Villacorta, Gilberto M.

; REGISTRATION NUMBER: 34,038

; REFERENCE/DOCKET NUMBER: 4980-007-0

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 413-3000

; TELEFAX: (703) 413-2220

; TELEX: 248855 OPAT UR


```
; INFORMATION FOR SEQ ID NO: 104:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 4 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 104:
US-09-938-315-104

Query Match          100.0%; Score 20; DB 9; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.1e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GSGS 4
      ||||
Db      1 GSGS 4

RESULT 2
US-09-949-039-54
; Sequence 54, Application US/09949039
; Publication No. US20030166160A1
; GENERAL INFORMATION:
; APPLICANT: HAWLEY, STEPHEN B.
; TITLE OF INVENTION: COMPOUNDS AND MOLECULAR COMPLEXES COMPRISING MULTIPLE
;   BINDING REGIONS DIRECTED TO TRANSCYTOTIC LIGANDS
; FILE REFERENCE: 057220/1301
; CURRENT APPLICATION NUMBER: US/09/949,039
; CURRENT FILING DATE: 2001-09-06
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 54
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Gly-Ser linker
US-09-949-039-54

Query Match          100.0%; Score 20; DB 12; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.1e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GSGS 4
      ||||
Db      1 GSGS 4

RESULT 3
US-10-262-435-90
; Sequence 90, Application US/10262435
; Publication No. US20030166832A1
; GENERAL INFORMATION:
; APPLICANT: Goldstein, Gideon
; TITLE OF INVENTION: Methods and Compositions for Impairing
;   Multiplication of HIV-1
; NUMBER OF SEQUENCES: 124
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr., P.O. Box 457
; CITY: Spring House
; STATE: PA
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/262,435
; FILING DATE: 30-Sep-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/451,067
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 09/113,921
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: GGP2AUSA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 4 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: <Unknown>
```

```

;
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 4
; OTHER INFORMATION: /note= "Biocytin-amide is attached
; to Ser in position 4."
; SEQUENCE DESCRIPTION: SEQ ID NO: 91:
US-10-262-435-91

Query Match      100.0%; Score 20; DB 12; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.1e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GSGS 4
Db      1 GSGS 4

RESULT 5
US-10-323-013-39
; Sequence 39, Application US/10323013
; Publication No. US20030180326A1
; GENERAL INFORMATION:
; APPLICANT: Goldstein, Gideon
; TITLE OF INVENTION: Methods and Compositions for Impairing Multiplication of HIV-1
; FILE REFERENCE: GGP3USA
; CURRENT APPLICATION NUMBER: US/10/323,013
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: US/10/114,176
; PRIOR FILING DATE: 2002-04-02
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 39
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-10-323-013-39

Query Match      100.0%; Score 20; DB 12; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.1e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GSGS 4
Db      1 GSGS 4

RESULT 6
US-10-161-791-104
; Sequence 104, Application US/10161791
; Publication No. US2003018683A1
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: KAY, Brian K.
; APPLICANT: THORN, Judith M.
; APPLICANT: QUILLIAM, Lawrence A.
; APPLICANT: DER, Channing J.
; APPLICANT: FOWLES, Dana M.
; APPLICANT: RIDER, James E.
; TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
; TITLE OF INVENTION: ISOLATING AND USING SAME
; NUMBER OF SEQUENCES: 467
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

```

;
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/161,791
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/602,999
; FILING DATE: 16-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-202
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 104:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-10-161-791-104

Query Match      100.0%; Score 20; DB 12; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.1e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GSGS 4
Db      1 GSGS 4

RESULT 7
US-10-086-208-90
; Sequence 90, Application US/10086208
; Publication No. US20030194408A1
; GENERAL INFORMATION:
; APPLICANT: Goldstein, Gideon
; TITLE OF INVENTION: Methods and Compositions for Impairing
; Multiplication of HIV-1
; NUMBER OF SEQUENCES: 124
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr., P.O. Box 457
; CITY: Spring House
; STATE: PA
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/086,208
; FILING DATE: 28-Feb-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/451,067
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 09/113,921
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: GGP2AUSA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 90:
; SEQUENCE CHARACTERISTICS:

```

LENGTH: 4 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 90;
US-10-086-208-90

Query Match 100.0%; Score 20; DB 12; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.1e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSGS 4
DB 1 GSGS 4

RESULT 8
US-10-086-208-91
; Sequence 91, Application US/10086208
; Publication No. US20030194408A1
; GENERAL INFORMATION:
; APPLICANT: Goldstein, Gideon
; TITLE OF INVENTION: Methods and Compositions for Impairing
; Multiplication of HIV-1
; NUMBER OF SEQUENCES: 124
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr., P.O. Box 457
; CITY: Spring House
; STATE: PA
; COUNTRY: USA
; ZIP: 19477

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/086,208
FILING DATE: 28-Feb-2002
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/451,067
FILING DATE: <Unknown>

APPLICATION NUMBER: 09/113,921
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.

REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: GGP2AUSA

TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818

INFORMATION FOR SEQ ID NO: 91:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>

TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 4

OTHER INFORMATION: /note= "Biocytin-amide is attached
to Ser in position 4."
SEQUENCE DESCRIPTION: SEQ ID NO: 91;
US-10-086-208-91

Query Match 100.0%; Score 20; DB 12; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.1e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSGS 4
DB 1 GSGS 4

RESULT 9
US-10-044-844-8
; Sequence 8, Application US/10044844
; Publication No. US20040009904A1
; GENERAL INFORMATION:
; APPLICANT: Victoria, Edward J.
; Marquis, David M.
; Jones, David S.
; Yu, Lin
; TITLE OF INVENTION: aPL IMMUNOREACTIVE PEPTIDES, CONJUGATES THEREOF AND
; METHODS OF TREATMENT FOR aPL ANTIBODY-MEDIATED PATHOLOGIES
; NUMBER OF SEQUENCES: 226
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER LLP
; STREET: 755 PAGE MILL ROAD
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/044,844
FILING DATE: 10-Jan-2002
CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: CATHERINE M. POLIZZI
REGISTRATION NUMBER: 40,130
REFERENCE/DOCKET NUMBER: 25231-20061.03

TELEPHONE: (650) 813-5600
TELEFAX: (650) 494-0792
TELEX: 706141

INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 8;
US-10-044-844-8

Query Match 100.0%; Score 20; DB 12; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.1e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSGS 4
DB 1 GSGS 4

RESULT 10
US-10-114-176-39
; Sequence 39, Application US/10114176
; Publication No. US2002019232A1
; GENERAL INFORMATION:
; APPLICANT: Goldstein, Gideon
; TITLE OF INVENTION: Methods and Compositions for Impairing Multiplication of HIV-1

FILE REFERENCE: GGP3USA
CURRENT APPLICATION NUMBER: US/10/114,176
CURRENT FILING DATE: 2002-04-02

NUMBER OF SEQ ID NOS: 39
SOFTWARE: Patent in version 3.0
SEQ ID NO 39
LENGTH: 4

QY 1 GSGS 4
DB 1 GSGS 4

```
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-10-114-176-39

Query Match
Best Local Similarity 100.0%; Score 20; DB 14; Length 4;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSGS 4
Db 1 GSGS 4

RESULT 11
US-10-133-133A-7
; Sequence 7, Application US/10133133A
; Publication No. US20030114385A1
; GENERAL INFORMATION:
; APPLICANT: CATHERS, Brian
; APPLICANT: NEUTEBOON, Saskia
; APPLICANT: SHEPARD, Michael
; TITLE OF INVENTION: VIRAL ENZYME ACTIVATED PROTOXOPHORES
; TITLE OF INVENTION: AND USE OF SAME TO TREAT VIRAL INFECTIONS
; FILE REFERENCE: NB 2021.00
; CURRENT APPLICATION NUMBER: US/10/133,133A
; PRIOR FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: 60/286,983
; PRIOR FILING DATE: 2001-04-27
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Linker sequence between NS3 and NS4A polypeptides
US-10-133-133A-7

Query Match
Best Local Similarity 100.0%; Score 20; DB 15; Length 4;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSGS 4
Db 1 GSGS 4

RESULT 12
US-10-099-424B-8
; Sequence 8, Application US/10099424B
; Publication No. US20030100127A1
; GENERAL INFORMATION:
; APPLICANT: Corn, Robert M.
; APPLICANT: Smith, Emily
; APPLICANT: Weisblum, Bernard
; APPLICANT: Erickson, Matthew
; APPLICANT: Uljasz, Andrew
; APPLICANT: Wanat, Matthew
; TITLE OF INVENTION: Fusion Protein Arrays on Metal Substrates for Surface Plasmon Resonance
; TITLE OF INVENTION: Imaging
; FILE REFERENCE: 09820.210
; CURRENT APPLICATION NUMBER: US/10/099,424B
; PRIOR FILING DATE: 2002-03-15
; PRIOR FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: 5-residue polypeptide linker
```

```
US-10-099-424B-8

Query Match
Best Local Similarity 100.0%; Score 20; DB 15; Length 5;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSGS 4
Db 2 GSGS 5

RESULT 13
US-10-274-638-6
; Sequence 6, Application US/10274638
; Publication No. US20030109000A1
; GENERAL INFORMATION:
; APPLICANT: Moore, Margaret D.
; APPLICANT: Fox, Brian A.
; TITLE OF INVENTION: DIMERIZED GROWTH FACTOR AND MATERIALS
; TITLE OF INVENTION: AND METHODS FOR PRODUCING IT
; FILE REFERENCE: 01-30
; CURRENT APPLICATION NUMBER: US/10/274,638
; CURRENT FILING DATE: 2002-10-18
; PRIOR FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: polypeptide, linker peptide
US-10-274-638-6

Query Match
Best Local Similarity 100.0%; Score 20; DB 15; Length 5;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSGS 4
Db 2 GSGS 5

RESULT 14
US-10-412-382-17
; Sequence 17, Application US/10412382
; Publication No. US20040009507A1
; GENERAL INFORMATION:
; APPLICANT: Domantis, Ltd.
; APPLICANT: Winter, Gregory P.
; APPLICANT: Jespers, Laurent
; APPLICANT: Lasters, Ignace
; APPLICANT: Wang, Peter
; TITLE OF INVENTION: CONCATENATED NUCLEIC ACID SEQUENCES
; FILE REFERENCE: 8039/2062
; CURRENT APPLICATION NUMBER: US/10/412,382
; CURRENT FILING DATE: 2003-04-11
; PRIOR APPLICATION NUMBER: PCT/GB01/04615
; PRIOR FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: GB 0025144.7
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 17
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Linker
US-10-412-382-17

Query Match
Best Local Similarity 100.0%; Score 20; DB 12; Length 7;
```

Best Local Similarity 100.0%; Pred. No. 7.1e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GSGS 4
Db 3 GSGS 6

RESULT 15

US-10-104-919-48
; Sequence 48, Application US/10104919
; Publication No. US20030099608A1
; GENERAL INFORMATION:
; APPLICANT: Presnell, Scott R.
; APPLICANT: Xu, Wenfeng
; APPLICANT: Kindsvogel, Wayne
; APPLICANT: Chen, Zhi
; APPLICANT: Hughes, Steven D.
; TITLE OF INVENTION: Human Cytokine Receptor
; FILE REFERENCE: 01-12
; CURRENT APPLICATION NUMBER: US/10/104,919
; CURRENT FILING DATE: 2002-03-23
; PRIOR APPLICATION NUMBER: US 60/279,222
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 48
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Gly-Ser spacer peptide
US-10-104-919-48

Query Match 100.0%; Score 20; DB 15; Length 8;
Best Local Similarity 100.0%; Pred. No. 7.1e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GSGS 4
Db 1 GSGS 4

Search completed: February 11, 2004, 22:24:51
Job time : 78 secs